NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

A MULTIVARIATE ANALYSIS OF LOST WORK TIME DUE TO ON-THE-JOB INJURIES AT MARINE CORPS COMMANDS

by

Timothy J. Robinson

September 2007

Thesis Advisor:	Lyn R. Whitaker
Second Reader:	Sergio Posadas

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2007	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE A Multivariate Analysis of Lost Work Time Due to on-the-Job Injuries at Marine Corps Commands			**5. FUNDING NUMBERS**
6. AUTHOR(S) Timothy J. Robinson			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

Yearly, the Department of the Navy pays about $245 million in workers' compensation and related medical benefits under the Federal Employee Compensation Act program. (Bowes, 2003) Based on data from the Defense Manpower Data Center (DMDC) and Defense Finance and Accounting Services (DFAS), the Office of the Secretary of Defense stated that since fiscal year 2002, the number of lost workdays (LWD) the United States Marine Corps (USMC) has accumulated per hundred civilian employees has been higher than the rate for the United States Army (USA), United States Navy (USN) and United States Air Force (USAF).

This thesis investigates the LWD rate of the USN and the USMC, with more detailed analysis on the USMC. The goal is to identify factors that lead to a high LWD rate and to find out which employees are more likely to accrue LWD.

This study consists of the use of generalized additive models, classification trees, and descriptive statistics to explore historic datasets to determine which factors influence an employee's tendency to accrue a LWD the most.

It is found that fire fighters, mechanics and police followed by equipment operators under the GS10 pay grade are at greatest risk of accruing at least one LWD per year.

14. SUBJECT TERMS Lost Workday, Multivariate Linear Regression, Generalized Additive Models, Classification and Regression Trees			**15. NUMBER OF PAGES** 105
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A MULTIVARIATE ANALYSIS OF LOST WORK TIME DUE TO ON-THE-JOB INJURIES AT MARINE CORPS COMMANDS

Timothy J. Robinson
Captain, United States Marine Corps
B.S., Virginia Polytechnic Institute & State University, 2000

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2007

Author: Timothy J. Robinson

Approved by: Lyn R. Whitaker
Thesis Advisor

Sergio Posadas
Second Reader

James N. Eagle
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Yearly, the Department of the Navy pays about $245 million in workers' compensation and related medical benefits under the Federal Employee Compensation Act program. (Bowes, 2003) Based on data from the Defense Manpower Data Center (DMDC) and Defense Finance and Accounting Services (DFAS), the Office of the Secretary of Defense stated that since fiscal year 2002, the number of lost workdays (LWD) the United States Marine Corps (USMC) has accumulated per hundred civilian employees has been higher than the rate for the United States Army (USA), United States Navy (USN) and United States Air Force (USAF).

This thesis investigates the LWD rate of the USN and the USMC, with more detailed analysis on the USMC. The goal is to identify factors that lead to a high LWD rate and to find out which employees are more likely to accrue LWD.

This study consists of the use of generalized additive models, classification trees, and descriptive statistics to explore historic datasets to determine which factors influence an employee's tendency to accrue a LWD the most.

It is found that fire fighters, mechanics and police followed by equipment operators under the GS10 pay grade are at greatest risk of accruing at least one LWD per year.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First of all, I would like to thank God. I'm here because of His grace and mercy. I would like to thank my beautiful wife who has "weathered the storm" when it comes to being an Operations Research students' wife. She really has set the standard, in my mind, of what a wife really should be. I also want to acknowledge Professor Lyn Whitaker, the smartest professor I know, and Lieutenant Colonel Posadas for helping me climb this "thesis mountain." Last, but not least, I would like to acknowledge the "Sled Dogs" of our cohort. You know who you are.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Yearly, the Department of the Navy (DoN) pays about $245 million in workers' compensation and related medical benefits under the Federal Employee Compensation Act (FECA) program. (Bowes, 2003) Based on data from the Defense Manpower Data Center (DMDC) and Defense Finance and Accounting Services (DFAS), the Office of the Secretary of Defense (OSD) states that since fiscal year (FY) 2002, the number of lost workdays (LWD) (per hundred civilian employees per year) in the United States Marine Corps (USMC) has been higher than rate for the United States Army (USA), United States Navy (USN) and United States Air Force (USAF).

"In an attempt to reduce the costs of the DoN Workers' Compensation program and provide the employee with a safe working environment, the Secretary of the Navy (SECNAV) established a goal for the USN and the USMC to reduce their LWD by 70 percent and 85 percent, respectively, by 2006, from their annual level as of FY 2001." (Naval Audit Service [NAS], 2004). The USMC has not attained this goal. According to the NAS, by June 2003 the USMC LWD rate per hundred employees was 85.1 days. This measure is 70 days higher than the projected FY 2006 goal of 11.1 days per hundred employees.

This thesis investigates the LWD rate of the USMC. USN LWD rates are also studied for comparison. The goal is to identify factors that lead to a high LWD rate and to find out which employees are more likely to accrue LWD.

The first portion of this study consists of a careful investigation into how LWD are recorded, how LWD rates are computed, and the structure of the two major databases concerning LWD; the Defense Manpower Data Center DMDC data, and the Naval Safety Center NSC data. This portion of the study is based in part on site visits to MCLB Albany and MCLB Barstow. The second portion of the study focuses on the analysis of the DMDC data. We use both, descriptive statistics and classification trees to study the type of employees who have high

LWD rates. A binary (1 or 0) variable was assigned to all employees who accrued a LWD in order to separate those with or without LWD. Factors such as civilian occupation codes, age, pay grade, and type of installation were studied to determine which factors influenced an employee's tendency to accrue a LWD the most.

The USMC has not achieved the SECNAV's LWD rate goal. A possible reason for this is an increase in employee hiring. As the rate of adding new employees increases, the LWD totals tend to increase. Since the base year of 2002, the percent change in LWD totals have matched increases and decreases in the percent change in the total number of employees. Figure 1 shows the percentage increase or decrease in total LWD and employees based on the previous year.

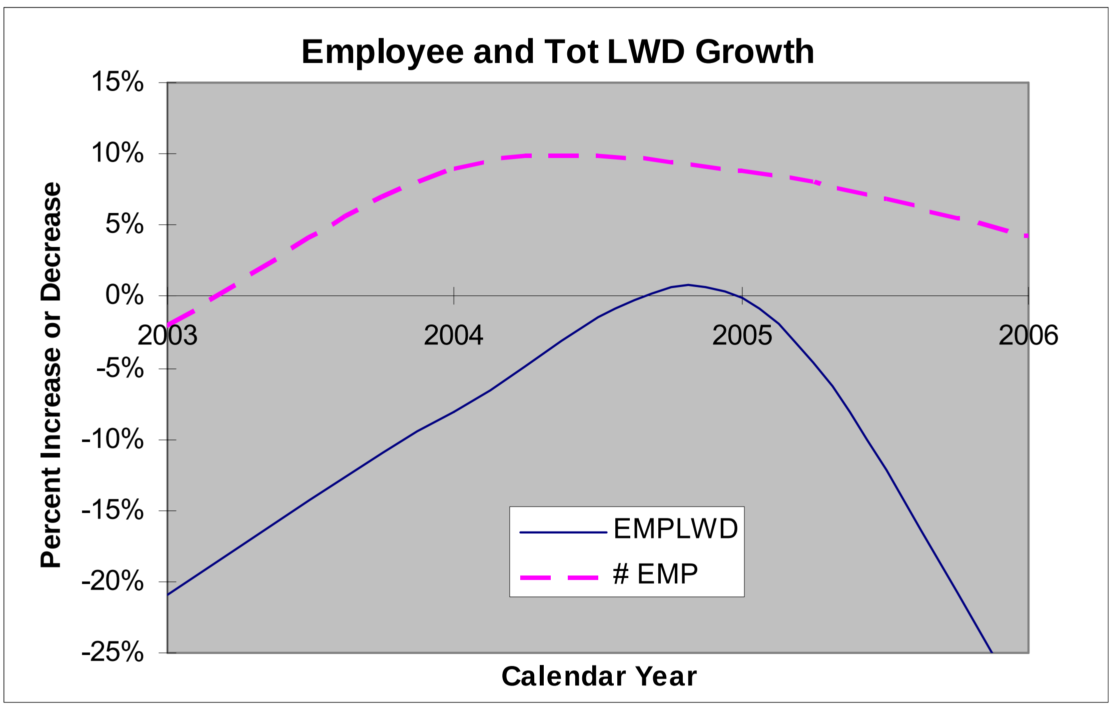


Figure E.1 CY vs. Percent change in employees and LWD

Since the base year of 2002, the percent change in LWD totals have mimicked increases and decreases in the percent change in the total number of employees. The graph shows that an increase in employees can be directly related to an increase in LWD. For example, throughout early 2003, the percent change in number of employees increased. Likewise the percent change in LWD totals increased in this same period. From 2005 to 2006 the percent change in the number of employees decreased along with the percent change in the number of LWD. However, an increase in total LWD does not necessarily imply an increase in the LWD rate. From 2003 to 2006, the civilian workforce has risen from 14750 to 17845, a total of 17.3%. Despite this significant increase, the USMC has decreased the annual LWD_{rate} by 10%.

The process of recording and computing LWD revealed a few potential difficulties such differences in LWD recording practices at MCLB Albany and Barstow and how DMDC identifies USMC civilian employees in their database. Whether these influence the LWD computations much is not known. Of greater issue is the difficulty in using the NSC data to study the details of the LWD cases. Several large discrepancies were found between the NSC and the DMDC data. Thus analysis focused on DMDC data.

Analysis of the DMDC data reveals that fire-fighters, security forces, and mechanics are at the greatest risk for accruing a LWD (Figure E.2). The next high risk group is equipment operators that are of pay grade GS10 or below. More detailed analysis reveals that differences in LWD rates can be in part accounted for by differences in the types of employees at each base.

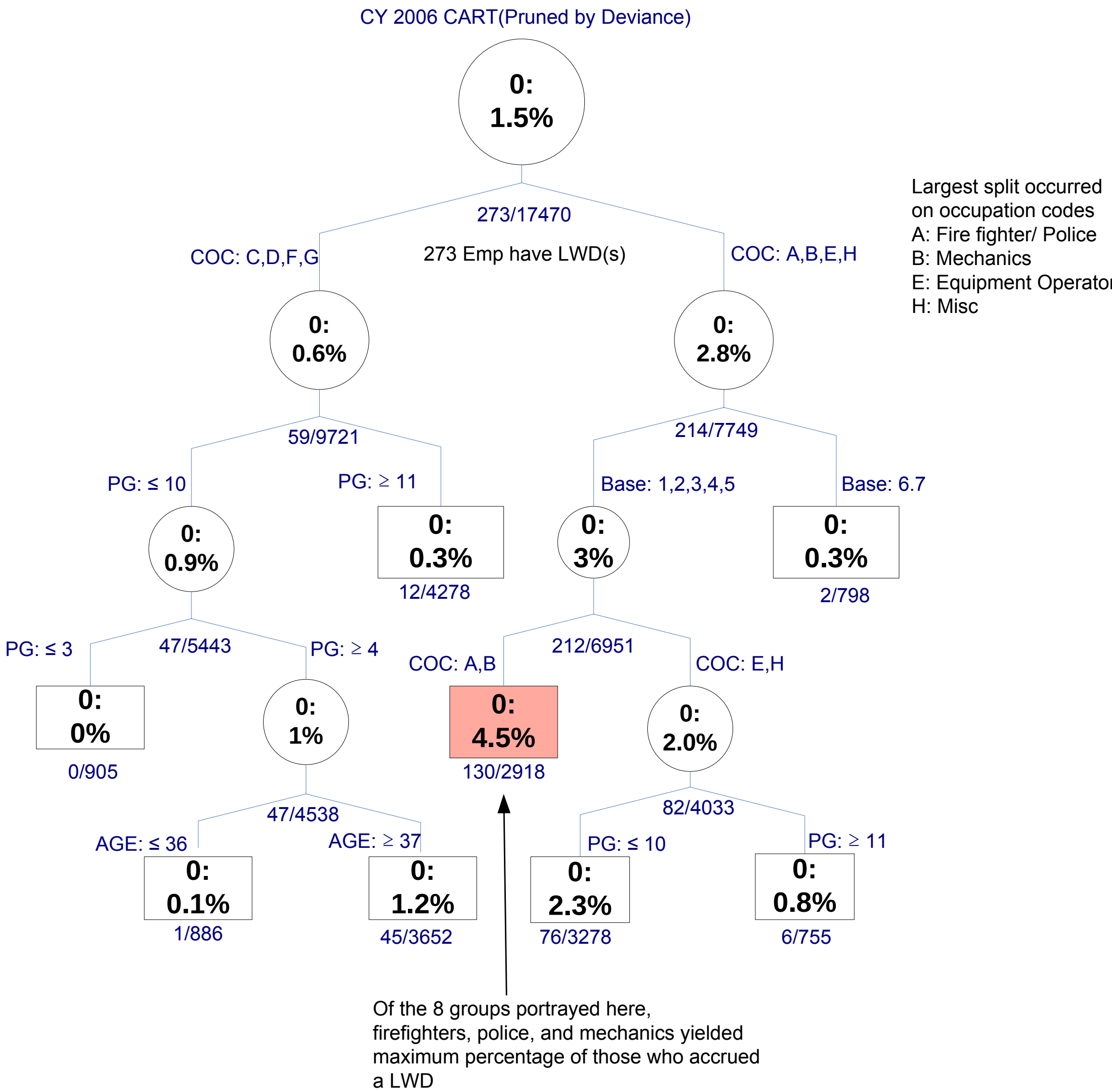


Figure E.2. CT USMC CY 2006

Personnel LWD rates for the entire population of USMC employees were compared to those from am equally-sized random sample of USN employees. This comparison was based on LWD rate per individual employee. The individual-employee LWD rates between the services are different. The greatest disparities can be found in those who have individual-employee LWD rates less than 0.16. This group consists of approximately 244 USMC employees and 121

USN employees (Figure 3). The USMC has a larger percentage of employees that incur LWD; however, USN employees produce more LWD per individual. The individual-employee LWD rate is reflected in the histogram below.

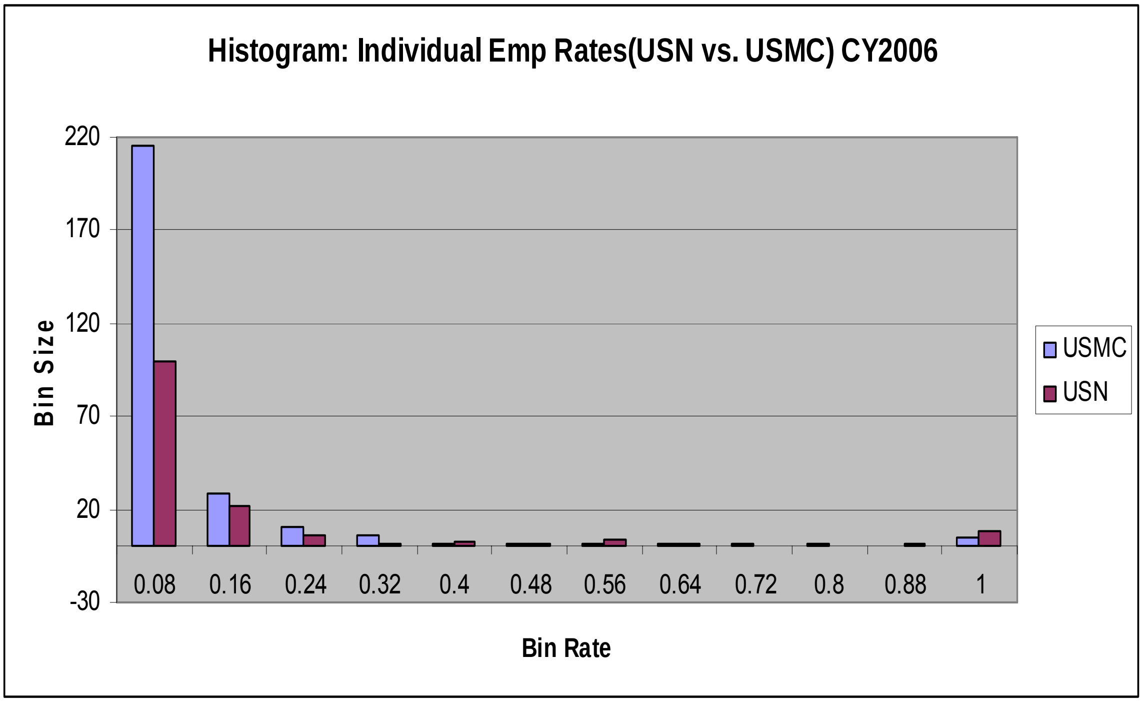


Figure E.3. Histogram of Individual Employee Rates CY 2006

Another notable find was that 44% of all LWD recorded for CY 2006 was from employees who were accruing LWD in CY 2005. It is assumed that these LWD were from prior injuries. Therefore, two metrics would better represent the scale of LWD: adjusted LWD rate (ALR) and focused LWD rate (FLR). ALR would focus only on new cases that accrued LWD within the current year whereas FLR would focus on the LWD rate of only those with at least one LWD. Cases that incurred a high number of LWD and carried over into the next CY are referred to as outliers. The LWD rate without these outliers for CY 2006 is 26.76 per 100 employees; still higher than the 11.1 LWD rate goal for FY 2006, but a significant 36% drop from FY 2002.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

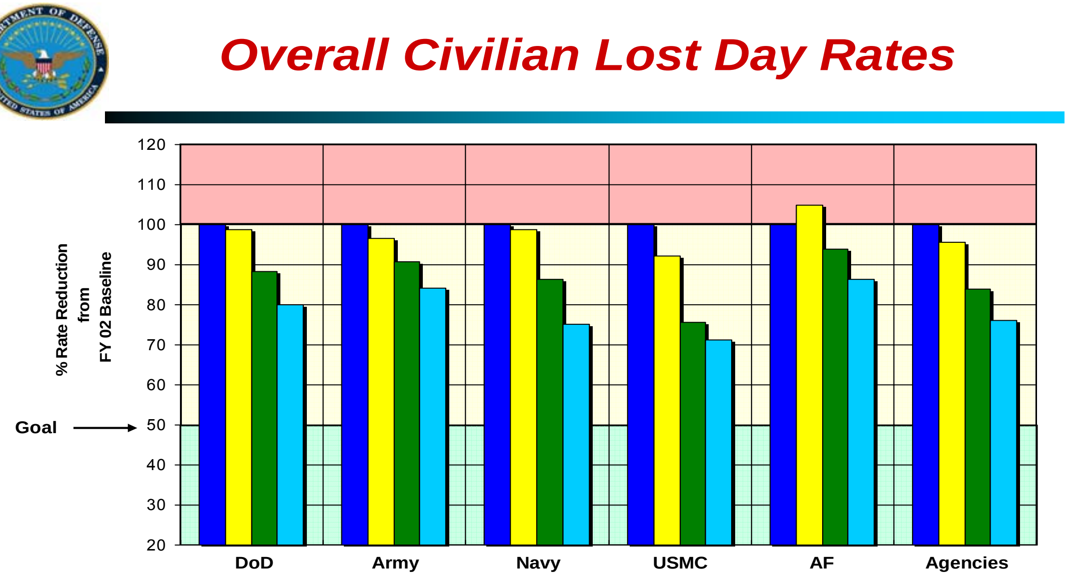
A. BACKGROUND

Yearly, the Department of the Navy (DoN) pays about $245 million in workers' compensation and related medical benefits under the Federal Employee Compensation Act (FECA) program. (Bowes, 2003) Based on data from the DMDC and DFAS, the Office of the Secretary of Defense (OSD) stated that since fiscal year (FY) 2002, the number of LWD the USMC has accumulated per 100 employees has been higher than the number of LWD accumulated by the USA, USN, and USAF. (See Figure 1.1)

"In an attempt to reduce the costs of the DoN Workers' Compensation program and provide the employee with a safe working environment, the Secretary of the Navy (SECNAV) established a goal for the USN and the USMC to reduce their LWD by 70 percent and 85 percent, respectively, by 2006, from their annual level as of FY 2001." (Naval Audit Service, 2004). The USMC has not attained this goal. According to the Naval Audit Service (NAS), by June 2003 the USMC lost-workday rate per 100 employees was 85.1 days. This measure is 70 days higher than the projected FY 2006 goal of 12.1 days per 100 employees.

This thesis investigates the USMC lost workday rates between the USN and the USMC and seeks to identify the variables or circumstances that lead to this difference.

Problems associated with computing the lost workday rate are multi-fold. It is not uncommon for reporting procedures to be problematic. For example, there is disparity in reporting procedures for USMC commands. (See Chapter III) Furthermore, there are different criteria for reporting LWD based on jobs specialties in the USMC. When comparing job specialties of the USN to the USMC, the two services have different missions and thus employ civilians in different ways. For example, both services may have heavy equipment operators, but the USN civilians may work in shipyards while the USMC civilians may work at a logistics base. These are two different environments with different risks that may lead to higher lost workday rates.



	DoD	Army	Navy	USMC	AF	Agencies
Baseline FY 02	31.49	29.29	39.83	73.83	25.60	25.58
FY 03	31.07	28.29	39.39	67.97	26.85	24.46
FY 04	27.86	26.34	34.74	55.67	24.16	21.46
FY 05	25.23	24.62	29.92	52.57	22.08	19.48
FY 05 Goal	15.74	14.65	19.92	36.92	12.80	12.79

Source: DMDC/DFAS pay records, days lost per 100 employees per year

OSD-1

Figure 1.1 Overall Civilian Lost Day Rates

Recent published studies suggest that a lack of uniformity in regional offices could be another factor in the extreme difference in LWD between services. According to Bowes, “the FECA program suffers from lack of central coordination…A more regionally coordinated approach to establishing job pools for return to work would be helpful. Oversight is too fragmented, with no one really aware of how individual FECA programs are performing.” (Bowes, 2003)

According to an audit the NAS performed, the number of LWD remained high “due to a lack of specific mandated requirements in current Navy and Marine Corps directives.” (NAS, 2004) The NAS also found a lack of consistent partnership between the Injury Compensation Program Administrators (ICPA), supervisors, and safety and occupational health departments. (NAS, 2004) Since distinct roles and responsibilities were not established between these entities, there is a possibility for error that may cause LWD rates to be artificially high or low.

B. OBJECTIVES AND LWD DATA

The purpose of this thesis research is to analyze why the lost workday rate for the USMC, as reported by the DMDC, is higher than the lost workday rate of the other services. This thesis investigates several factors that influence the USMC LWD rate. Furthermore, this thesis explores better ways to calculate and track the LWD rates.

This study includes a report on two site visits to MLCB Albany and MCLB Barstow to investigate how LWD data is reported prior to being sent to the DMDC and the Naval Safety Center (NSC). Analysis is based on data from the DMDC. The DMDC provided seven years of data, dating from calendar year (CY) 2000 to 2006. The NSC provided seven years of case data, dating from FY 2000 to 2006. Differences between the two datasets are explored in detail in Chapter III.

The focus of this study is on USMC CY 2006 number of LWD per hundred employees. The USMC as a whole is analyzed to find patterns or localized problems by base and type of job. USN data are analyzed for comparison purposes as well. The data to be analyzed will consist of all full-time and part-time, General Schedule Department of Defense (DoD) employees who were paid annually. Demographic variables such as Civilian Occupation Code (COC), COC Type, age, pay grade, command and command type are included.

It is assumed that since data from the DMDC is matched to data from the DFAS, which is responsible for payments to employees, the data is correct for the analytical purposes of this thesis. However, there is no validation process to reconcile differences between the DMDC and the NSC data. Therefore, when differences exist, the DMDC data is assumed to be more accurate. The link between the DFAS payments and the DMDC data drive the error correction in the DMDC data. With the NSC, if the amount of LWD logged for a particular person is incorrect, the employee's pay is not affected. Therefore all analysis is conducted using DMDC data.

C. THESIS ORGANIZATION

This thesis is organized into five chapters:

Chapter II: Definitions of key terms such as FECA, Continuation of Pay (COP), Leave Without Pay (LWOP), and Office of Workers' Compensation (OWCP) are given to understand how they relate to the problem. Three recent studies performed regarding LWD are also reviewed. The organizations that performed these studies include the Naval Audit Service (NAS), The Center for Naval Analysis (CNA), and the NSC. Additional personnel interviews and meetings were conducted during site visits to Marine Corps Logistics Base (MCLB), Albany and MCLB, Barstow.

Chapter III: This chapter discusses the data sources and the electronic lost workday data that was collected. It begins with the results of site visits to MCLB Albany and MCLB Barstow to document the LWD reporting practices at two large bases with similar missions. The data from DMDC and NSC databases are discussed. Data from theses sources are different because each organization has different missions. Moreover, each dataset must be processed differently. This chapter addresses data validation, database management, and potential problems with data as well.

Chapter IV: Using information/data received, an analysis is conducted. Using Classification Trees (CT) validated by a generalized additive model and descriptive statistics, conclusions are drawn.

Chapter V: This chapter gives a summary of findings and suggestions for further study and future research.

II. LITERATURE REVIEW

The analysis of LWD requires a careful definition of several key terms. In this chapter key terms are defined that will support understanding LWD. In addition, recent published studies on LWD performed by the NAS, the CNA, and the NSC are discussed in this chapter.

A. DEFINITIONS

1. FECA

FECA is administered by the Office for Worker's Compensation Programs (OWCP). It's mission is to provide "workers' compensation coverage to three million Federal and Postal workers around the world for employment-related injuries and occupational diseases." (Employment Standards Administration (ESA), 2007) The FECA provides benefits for medical and surgical care, rehabilitation, and compensation for wage replacement. With twelve district offices nationwide (see Figure 2.1), the FECA also is responsible for processing new claims, managing pre-existing claims, and assisting employees return to work after they are medically cleared. Lastly, in the untimely event of a work related injury or disease causing the death of and employee, the FECA provides for payment of benefits to dependents.

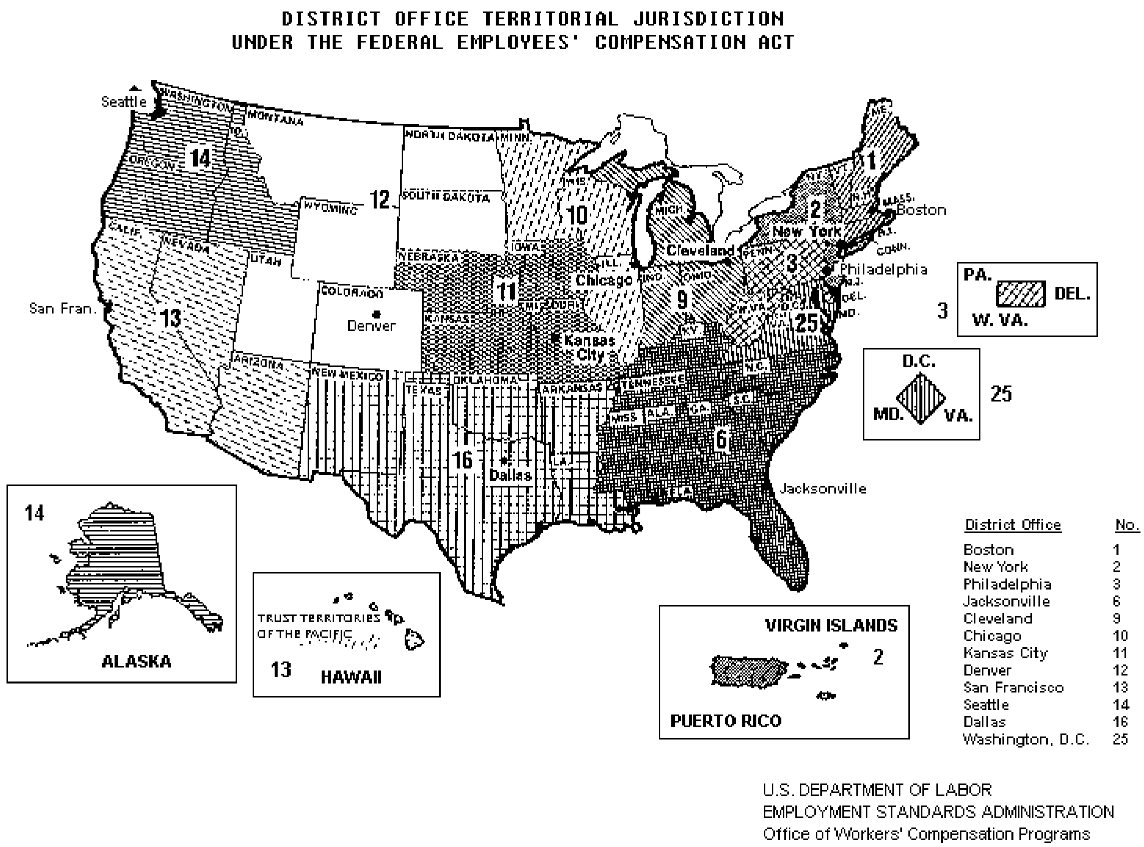


Figure 2.1 District Office Territorial Jurisdiction under FECA

2. COP

Employee entitlements depend on the number of days a federal employee is temporarily disabled due to an employment related injury or disease. If the employee is temporarily disabled from 0 to 45 days, he is entitled to COP. In this case the employing agency is responsible for compensation only up to 45 days. “If the employee has no dependents, compensation is generally payable at the rate of two-thirds of pre-disability gross wages tax-free. If the employee has one or more dependants, compensation is payable at the rate of three-fourths of pre-disability gross wages, tax-free." (ESA, 2007)[1] If the disability extends beyond the initial 45 days, the employee goes into a "non-paid status" called LWOP.

3. OWCP and LWOP

Three days after entering the non-paid status a federal employee is compensated for lost wages from the OWCP. The OWCP is responsible for administering the FECA, therefore they have similar functions and goals:

> OWCP seeks to protect the interests of eligible workers, employers and the Federal Government by ensuring timely and accurate claims adjudication and provision of benefits, by responsibly administering the funds authorized for this purpose, and by restoring injured workers to gainful work when permitted by the effects of the injury." (ESA, 2007)[2]

To cover wages for the three day non-paid waiting period, injured employees have the option of using their sick leave. Since employees get paid for sick leave, there will not be a gap in pay. The only time a waiting period is not required, is when the disability causing the wage loss lasts longer than 14 days from the time initial compensation begins. The compensation rates are the same for LWOP as they are for COP.

B. PREVIOUS LOST WORKDAY STUDIES

Only a few published studies have been conducted recently on LWD rates. While most are not USMC specific, they describe some of the underlying issues involved with LWD. What is consistent in the studies is the lingering effect that previous years' performance has had on current year LWD rates. For example, most studies indicate there is a problem with computation of accurate LWD rates or acknowledge the existence of some factor that attributed to the current calculated LWD rates. There are three published studies/analyses used in this thesis:

---

[1] Employment Standards Administration: U.S. Department of Labor, <http://www.dol.gov/esa/regs/compliance/owcp/91-18.htm>, 11 August 2007.

[2] Employment Standards Administration: U.S. Department of Labor, <http://www.dol.gov/esa/aboutesa/mission/owcp/owcpmiss.htm>, 11 August 2007.

1. Naval Audit Service Study

NAS published an audit entitled "Reducing Lost Work Time due to On-the-Job Injuries at Navy and Marine Corps Commands" on 26 March 2004. The objectives were:

> …to evaluate DoN's progress towards attaining lost-workday rate reduction goals set by SECNAV [Secretary of the Navy] and to evaluate the status of steps outlined by SECNAV to be taken by Navy and Marine Corps commanders to achieve the reduction. (NAS, 2004)

The SECNAV's goal was to reduce the FY 2001 COP days by 70% and FY 2001 LWOP days by 85% by FY 2006. By 2004 DoN commands did not show much progress towards achieving the SECNAV's LWD rate goals. Moreover, this was a good indication that DoN commands would not achieve the SECNAV's goals FY 2006.

There were several reasons for the slow progress achieved by DoN commands:

- The initial finding was a lack of specific mandated requirements in current USN and USMC directives. NAS also found an "overall lack of consistent partnerships between the ICPAs, supervisors, and safety and occupational health departments, and an overall lack of implementation of best practices and controls to reduce lost time due to injuries." (NAS, 2004)
- NAS found that some commands did not have active FECA Working Groups that analyzed historical data, established goals and identify strategies to achieve these goals with the abilities to hold managers and supervisors accountable. NAS also found that some commands had FECA working groups, but they excluded key personnel, such as the commanding officers, department heads, human resource personnel, and shop supervisors. (NAS, 2004)
- One of the most significant findings was that the DoN activities did not calculate or validate LWD rates. Since commands were not required to do so, they did not set up procedures to keep track of the number of COP and LWOP days their employees accumulated. Consequently, there is no way to be sure rates published by the DMDC are correct.

As of March 2004, the NAS position is USN and USMC commands have not implemented the best management practices in managing its worker's compensation program. The disparity in performance between DoN commands is realized due to a lack of central authority in monitoring progress or implementing best management practices in their FECA programs. According to the NAS, visibility and standardization of FECA programs across DoN would standardize reduce the LWD metric for the USN and USMC.

2. Center for Naval Analysis Study

"An Analysis of DoN Mishap Rates" is a study that was conducted in March 2004 by Michael D. Bowes for the CNA. The objective of this study is to "explore the improvement in mishap rates that occurred in the 1990's, determine the variety of factors that explain the decline in mishap rates, and develop a sound basis for projecting possible future reductions in mishap rates. (Bowes, 2004)

To execute this objective, Bowes looked at past factors that may have driven trends in the DoN civilian mishap rates. He also assessed and identified the differences in safety programs and evaluated the programs' effectiveness. Bowes recognized that sometimes rates will improve by other factors, such as unit safety efforts. Therefore, he analyzed the extent to which past improvements in mishap rates could be attributed to unit safety efforts and adjusted his conclusions accordingly. (Bowes, 2004)

Bowes' study revealed that approximately 50% of improvements in DoN mishap rates from 1990 – 1999 can be explained by demographic factors. This was largely in part to a declination in the size of the DoN's industrial workforce. The study revealed that safety improvements led to 2,600 fewer mishaps in 1998. The lower mishap rate led to a lower LWD rate. The study also found different LWD rates in similar job fields and activities across the DoN. Bowes concluded:

> We estimate that a 56-percent reduction in overall mishaps could be achieved if lagging activities improved enough to match the best-in-class performance among similar activities. The 56-percent reduction in annual mishaps results in lifetime savings of $81 million in workers' compensation costs. (Bowes, 2004)

Clearly one of the most effective ways to improve LWD rate is to reduce mishaps across commands. The added advantage is less money spent in worker's compensation costs.

3. Marine Corps Logistics Base (MCLB) Barstow Study

A study was conducted by the NSC entitled "Analysis of Lost Work Day Rate For Marine Corps Logistics Base Barstow, CA" dated September 26, 2006. NSC's task was to:

> …determine why MCLB Barstow's civilian lost time rate due to work-related injuries is so high and to provide a template to be used by Navy Echelon II and USMC Major Commands to enable them to perform similar analysis as a first step to improvement and as input to a Safety Center Analysis of all Don "Top 40" activities. (NSC, 2004)

The NSC found that employees elect to take LWOP instead of annual/sick leave which adds to the LWD metric. The study found that if overtime hours are taken out, LWD rates are barely affected. This is of particular importance due to the way the DMDC computes LWD rates which is discussed in Chapter III.

One of the most interesting findings was the accumulation of LWD for FY 2006 due to injuries sustained prior to FY 2006. According to the NSC, 88% of MCLB Barstow's FY 2006 LWD are attributed to injuries that occurred prior to FY 2006. Multiple data formats made consolidation of LWD information problematic. Data observed were from downloaded DMDC Lost Data Rate for USMC Installations, Occupational Safety and Health Administration (OSHA) 300 Logs for Base Ops provided to Barstow, and downloaded Web Enabled Safety System (WESS) Mishap Reports for LOGCOM.

In summary, NSC has noted continued reductions in the LWD rate of MCLB Albany. LWD resulting from previous injuries can account for a significant amount of LWD in the current year. Including data from previous injuries drives the LWD metric higher for the current year and is not an accurate depiction of the current year performance.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA AND LWD COMPUTATION

In this chapter the datasets used and the organizations that provide the data (DMDC and NSC) will be explained in detail. Also, reporting procedures from the organizational standpoint are explained. An explanation of how both datasets were converted and subsequently imported into several software applications for analysis is also given. Finally, descriptive statistics of both datasets are listed in tabular format.

The primary variable measured in this study is LWD. LWD can be expressed per person, per base or as a rate per 100 employees. All are defined, including the process of reporting LWD for MCLB Albany and MCLB Barstow. The two bases are modeled with a data flow diagram (DFD). There are minor subtleties that indicate differences in reporting procedures.

A. DATA REPORTING PROCEDURES

There is no standing operating procedure (SOP) that defines one overall method of collecting LWD data across USMC commands. The software applications and reporting requirements are the same; however, there is disparity in the way different USMC commands report LWD. On recent site visits to MCLB Albany and MCLB Barstow, this was found to be true. These two bases were chosen for several reasons. They both were found on the DMDC's Top 40 List numerous times. The DMDC's Top 40 List, lists DoD bases that have the highest LWD rate per 100 employees in descending order. The LWD rate is explained in Section C of this chapter. Another reason for choosing these bases is both provide multi-commodity depot level maintenance that support logistic operations for the USMC.

There are two basic reporting entities on both bases; the maintenance center and the base proper. The maintenance center and base proper do not use the same software applications for reporting. Before the reporting procedures for these two entities are described, it is important to understand some software applications that enable reporting.

1. Defense Civilian Personnel Data System (DCPDS)

The DCPDS is a web based human resources information system that supports civilian personnel operations in the DoD. (Civilian Personnel Management Service (CPMS), 2007)[3] The impetus behind DCPDS is to allow the DoD to move from multiple electronic systems to a single electronic management information system for DoD civilian employees.

According to the CPMS, there are three main reasons for DCPDS. The first reason is to reduce multiple systems. "In the past the DoD Components developed and used multiple and duplicative human resources information systems to accomplish the same type of work." (CPMS, 2007) DCPDS is a single system, capable of tying all DoD components together in an efficient, streamlined manner. The second reason for DCPDS is to reduce cost. Discontinuation of multiple systems with several operating requirements ultimately reduces cost. The third reason is to support regionalization. "The DCPDS will meet the dynamic needs of the Department with fewer personnel lists while responding to increased needs for HR [Human Resources] information." (CPMS, 2007)

Figure 3.1 is an illustration of the DCPDS operating environment. A data network connects information flow between Customer Support Units (CSU) and Regional Service Centers (RSC). In the case of MCLB Barstow and MCLB Albany, the CSU is the Budgeting or Fiscal Office. Employees and their

[3] Department of Defense, Civilian Personnel Management Service. <http://www.cpms.osd.mil/hrbits/userguide/dcpds_userguide.aspx>, 11 August 2007.

managers do not have access to DCPDS. The RSC is responsible for administration of data for all employees serviced by a specific region.

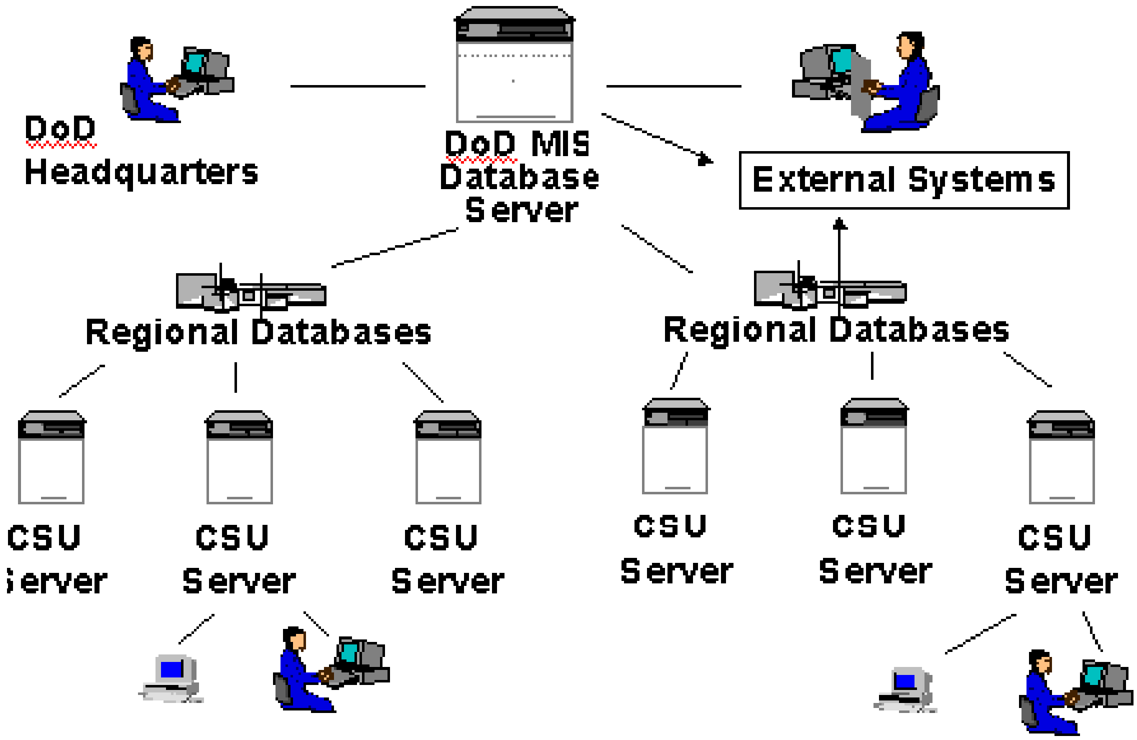


Figure 3.1 Illustration of the DCPDS Operating Environment

2. Standard labor Data Collection and Distribution Application (SLDCADA)

The SLDCADA (version 22.0) is a web-based application that is designed to be a timekeeping system that tracks DoD civilian, military, and contractor labor hours against Job Order Numbers (JONs) for financial and employee pay purposes. With the SLDCADA, Naval organizations enter and track time and attendance, check leave availability, make prior pay adjustments, input exceptions, and query the DCPDS. Perhaps the largest advantage of the SLDCADA is that it provides a single Time and Attendance screen for input, correction, certification, prior pays, and review (See Figure 3.2), which ultimately reduces the training effort and streamlines user/system interaction. (SLDCADA, 2007)

Figure 3.2 is an example of a biweekly time sheet interface for the SLDCADA. The SLDCADA can be accessed by employees for data entry. If there is a case where the employee cannot input his time or events, an administrator can update the SLDCADA on his behalf. The administrator is usually the supervisor of the shop or section.

SLDCADA BI-WEEKLY TIME AND ATTENDANCE SHEET

	Shop	Supv Assign	Name	SDA ID	Activity	Pay Period Date
Wk1	DOC01	DOCUMENT	DOCUMENTMIL JANE	SLDQA1	068561	05/07/2005
Wk2	DOC01	DOCUMENT	DOCUMENTMIL JANE	SLDQA1	068561	05/14/2005

Emp ID (SSN): 122122122 Type: Military FLSA: Emp Type: REGULAR

JON1: 1212
HRS: AWS: 0
JON2: 1272
HRS: Sched: SS
JON3:
HRS: Graded: G
Time &Ltr Cat: 5

	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Sched Wk1:		8.0	8.0	8.0	8.0	8.0		40.0
Wk2:		8.0	8.0	8.0	8.0	8.0		40.0
ND Wk1:								
Wk2:								
Shifts Wk1:	00	00	00	00	00	00	00	
Wk2:	00	00	00	00	00	00	00	

Week 1 Job Order Number	Per Shop / FCN Code	Env Haz	ND	THC	Sun	Mon	Tue	Wed	Thu	Fri	Sat

Week2 Job Order Number	Per Shop / FCN Code	Env Haz	ND	THC	Sun	Mon	Tue	Wed	Thu	Fri	Sat

Total:

Reg ____ OT ____ Comp ____ Hol ____ LV ____ Sun ____ 2nd ____ 3rd ____ ND ____ E/H ____ NP/LV ____

	Week1						Week2						
	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	Emp Int
Sun													
Mon													
Tue													
Wed													
Thu													
Fri													
Sat													

Employee Initial for Leave Taken. Note: If "LT"/"LU" is used, provide date of injury here (mm/dd/yyyy):

Available Injury Dates: ______________

CERTIFICATION: ATTENDANCE AND ABSENCES INCLUDING CHANGES MARKED HEREON ARE CERTIFIED CORRECT. LEAVE AND/OR OVERTIME APPROVED IN ACCORDANCE WITH EXISTING LAWS AND REGULATIONS.

I DID NOT SUFFER OR PERMIT ANY OVERTIME WORK OTHER THAN AS REPORTED HEREON FOR THIS WEEK, IF THIS EMPLOYEE IS NONEXEMPT.

Authorized Signature Date Telephone No.

This report may contain data subject to the Privacy Act of 1974.

Figure 3.2 Example of Biweekly Time and Attendance Sheet

3. Essex Replacement Program (ERP)

According to John Reed, the Systems and Financial Branch Manager at MCLB, Albany Maintenance Center, the former Essex program was a the data input system that read key punched cards into the Marine Corps Depot Maintenance Management System (DMMS). "DMMS was a batch processing system that the maintenance depots used for production control and Job Order Costing. The legacy Essex system with the DMMS, interfaced with the former Marine Corps Industrial Fund (MCIF) accounting system, now referred to as Navy Working Capital Fund (NWCF).

In the early 1990s the USMC modernized DMMS with a client server based system developed by Thompson Ramo Wooldridge Inc. (later renamed as TRW Inc.) that eventually replaced Essex. With no formal designation, the system created by TRW, became known as the "Essex Replacement Program". With this ERP system, data was punched in on time cards and manually keyed into DCPDS which meant dual data entry for job costing labor collection and payroll collection.

To alleviate multiple data entry processes and to standardize maintenance depot accounting systems Navair Industrial Financial Management System (NIFMS) was selected. "Instead of interfacing DMMS with NIFMS, the USMC retired DMMS, and implemented DIFMS/NIMMS [NAVAIR Industrial Material Management System] and a redesigned ERP."(Reed, 2007) Finally in 1998, the USMC Depots implemented a Manufacturing Resource Planning System (MRPII) that interfaces with ERP. ERP was modified again to include a sub-system that collected real-time artisan labor and completions at the work order level, creating an effective tracking and monitoring system for depot level maintenance.

4. Maintenance Center Data Flow

Figure 3.3 is a data flow diagram that depicts how MCLB Albany maintenance center reports. The sequence of events:

- The employee sustains an injury. It does not matter if the employee is suspected at fault; the injury must be reported.
- The employee can decide if he wants to take personal leave or elect COP[4].
- The employee's information is entered into the ERP by a supervisor.
- Before the end of the pay period the work the section supervisor prints out and audits all ERP records related to employee pay.
- At the end of the pay period ERP data is transferred into DCPDS via magnetic tape.
- Once data is into DCPDS, the budgeting office can make corrections up to a certain time period before data is processed by the DFAS and subsequently to the DMDC.

[4] Some situations the employee may be at fault. For example, if an employee sustained an injury due to his/her own negligence, the employee may take responsibility and use his/her own leave days to recover from an injury.

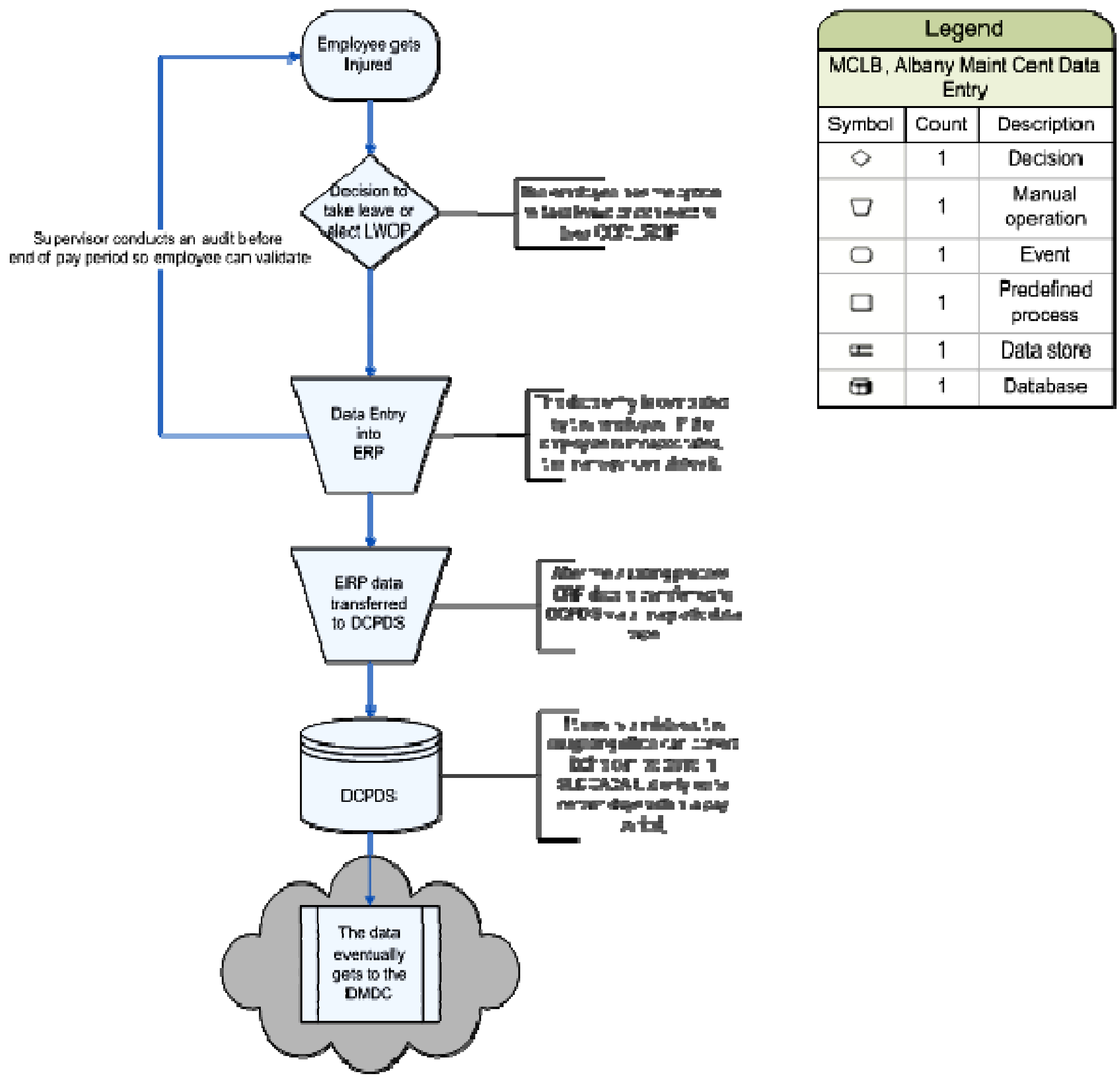


Figure 3.3 Data Flow Diagram: MCLB Albany Maintenance Center

Figure 3.4 is a data flow diagram that depicts how MCLB, Barstow maintenance center reports. The only difference is the way data is transferred from ERP to DCPDS. At MCLB Barstow, data is transferred electronically instead of manually from ERP to DCPDS by an automated script.

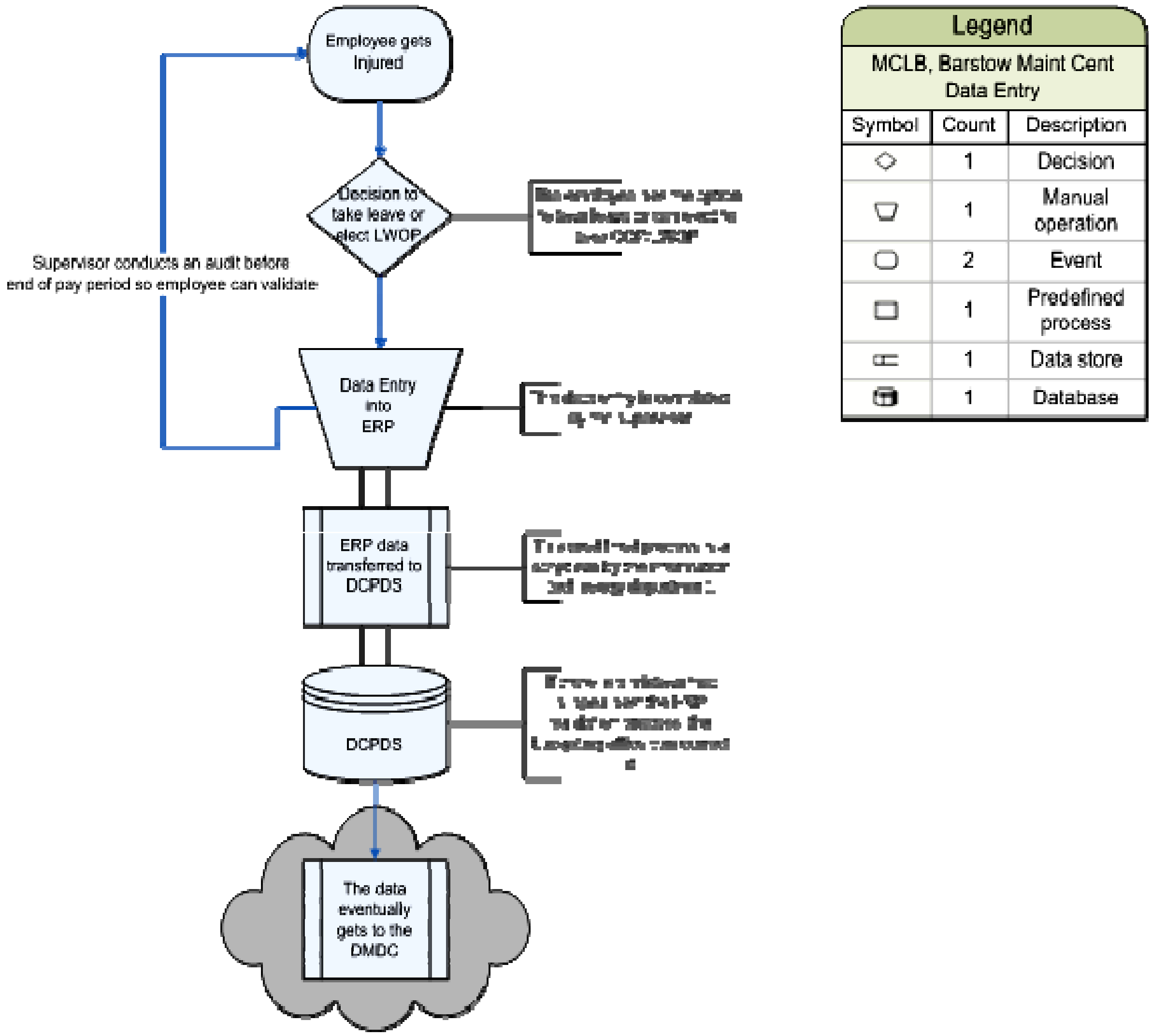


Figure 3.4 Data Flow Diagram: MCLB Barstow Maintenance Center

5. MCLB proper data flow

Figure 3.5 is a data flow diagram that depicts how MCLB, Albany base proper reports. The sequence of events is:

- The employee sustains an injury. It does not matter if the employee is suspected at fault; the injury must be reported.
- The employee can decide if he wants to take personal leave or elect COP/LWOP.

- The employee input's his/her information directly into SLDCADA. If the employee is incapacitated, his supervisor makes the necessary data entries.
- Before the end of the pay period the work section supervisor prints out and audits allows the employees to validate the data entry before final submission.
- Since SLDCADA is interfaced with DCPDS, there is a predefined process that transfers the data.

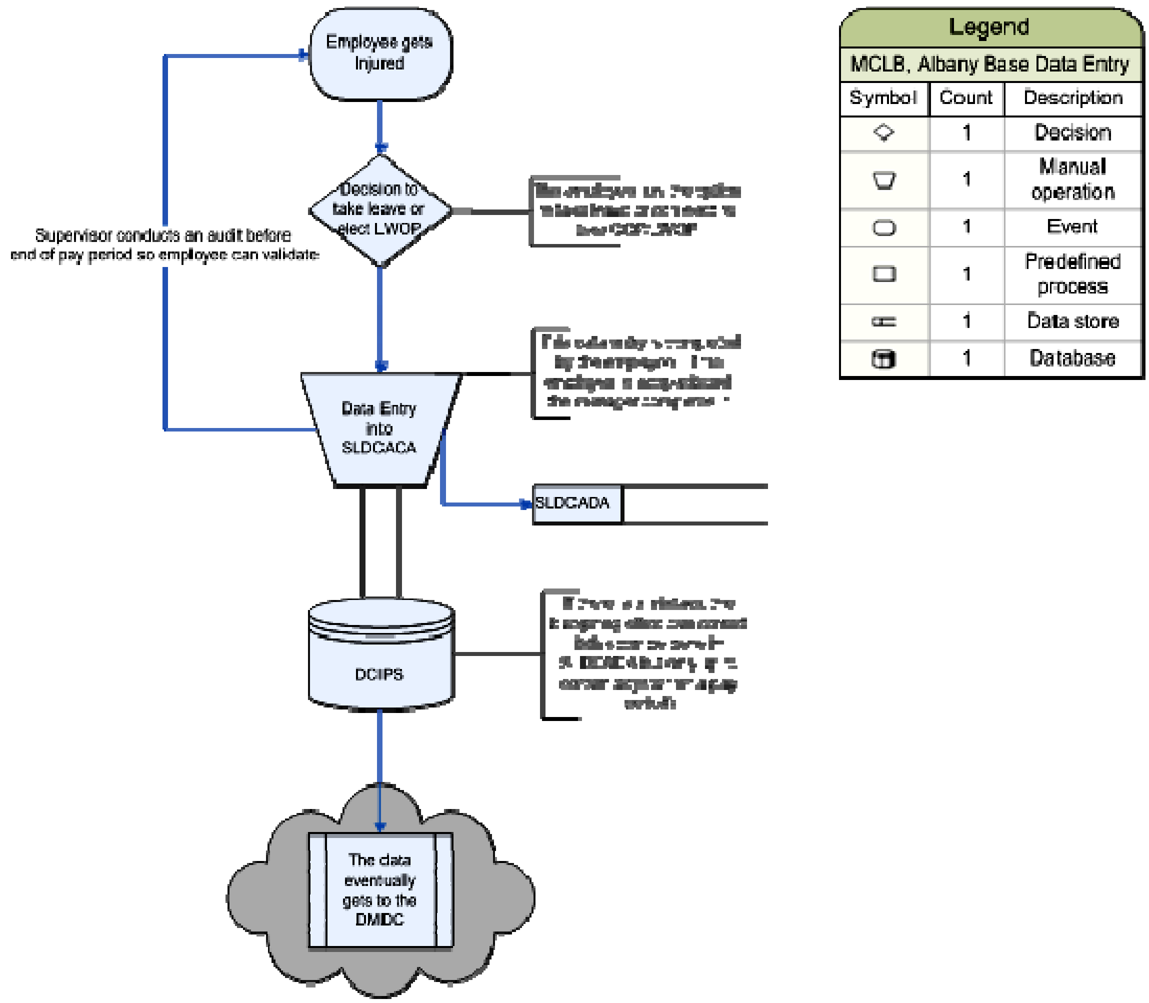


Figure 3.5 Data Flow Diagram: MCLB Albany Base

There are significant differences between MCLB Albany and Barstow base proper. Figure 3.6 is a data flow diagram that depicts how MCLB Barstow (base proper) reports. The most significant difference is the way in which data is stored and loaded into SLDCADA. MCLB Barstow stores weekly employee data on timesheets. Only section supervisors are able to enter data onto the timesheets. Before the end of the pay period these timesheets are reviewed by a department authorizing official for accuracy and then transferred to the Base Budgeting Office. Finally, the budgeting office has employees who enter the data into SLDCADA. If there is a mistake in SLDCADA, they can be corrected in SLDCADA or DCPDS before the end of the pay period.

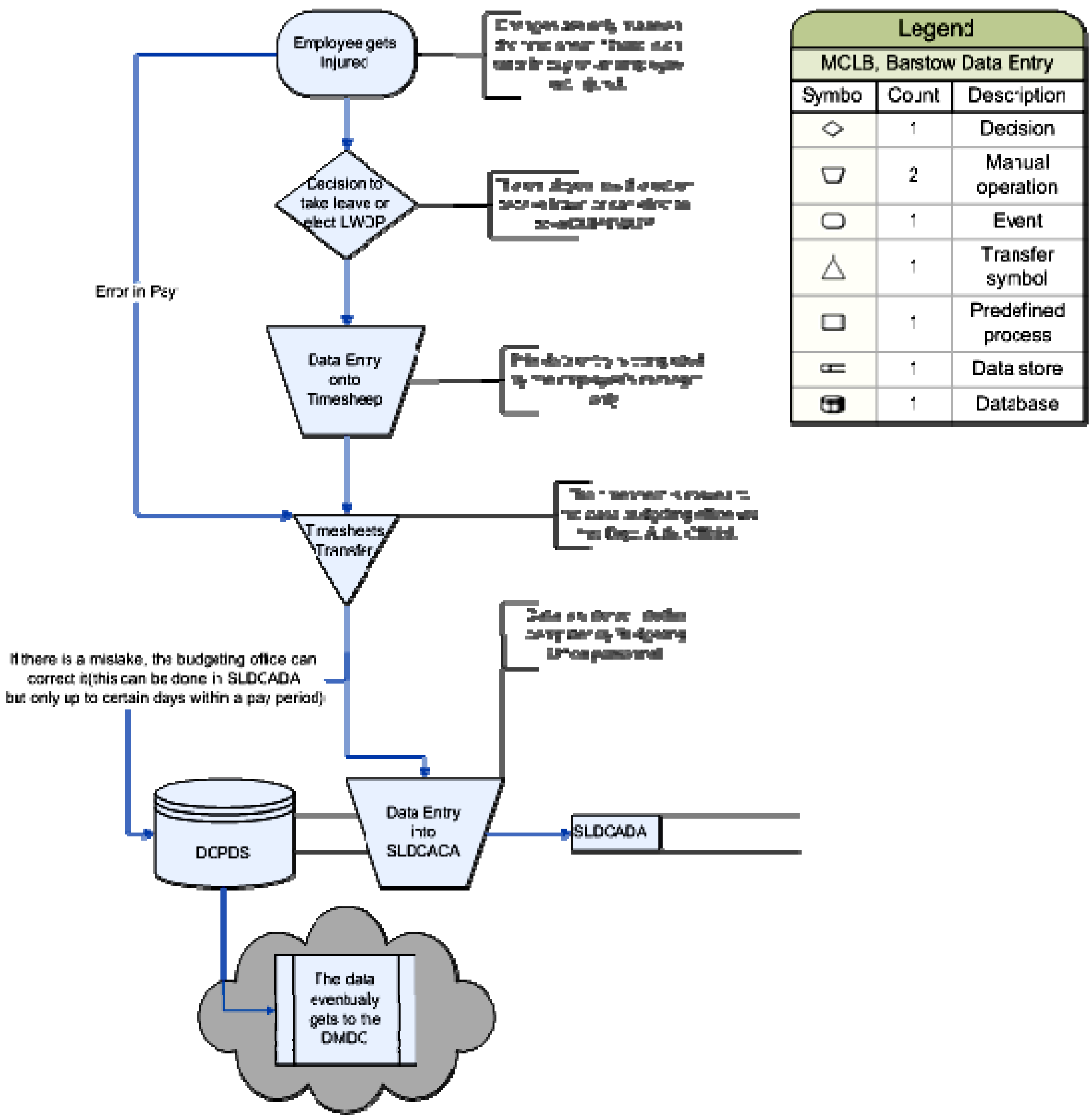


Figure 3.6 Data Flow Diagram: MCLB Barstow Base

In summary, there are slight differences in the reporting procedures between bases and within activities on certain bases. The extent of these differences may or may not affect the total LWD rates, but need to be taken into consideration.

B. LOST WORKDAY DEFINED

This definition of LWD is set forth by the DMDC. (DMDC, 2007) For a particular year, the number of LWD can be expressed in the summation represented in Equation (3.1). Index i represents a specific employee and index j represents a specific pay period (1 to 25 or 26, depending on the year). The LWD variable is normally considered an integer value, but due to partial workdays it is generally expressed as a fractional value.

$$LWD = \sum_{ij} (COP + LWOP)_{ij} \qquad (3.1)$$

The primary measure of performance (MOP) for the DMDC is LWD rate per 100 employees. Equation (3.2) represents how the DMDC calculates the LWD rate per 100 employees.

$$LWD_{rate} = \frac{\sum_{ij} (COP + LWOP)_{ij}}{\sum_{ij} (\text{prorated days})_{ij}} \times 26{,}000^{*} \qquad (3.2)$$

Note:
*The DMDC website uses 200,000 for a scalling factor derived from ($100_{employees} \times 40_{hours} \times 50_{weeks}$). A scaling factor of 26,000 is used for this thesis:
($100_{employees} \times 10_{days} \times 26_{pay\ periods}$), (DMDC, 2007)

Prorated days consist of the sum of regular (RegDays) worked, sick leave, annual leave, COP days, and LWOP days. For reporting purposes the DMDC computes rates by service, base, and major command.

C. DATASETS

The data pertaining to LWD comes primarily from two data sources, the DMDC located in Seaside, California and the NSC located in Norfolk, Virginia. Both organizations provided data for CY(s) 2000 – 2006 with all privacy act[5] information such as name, social security number (SSN), and address, etc. removed. To match demographics data to LWD, DMDC did provide a unique identifier for each individual.

There are several reasons for requesting data from the DMDC and NSC. The DMDC publishes the Top 40 list. The Top 40 list ranks, by LWD rate per hundred employees, all DoD bases and services in descending order. In order to answer several questions regarding the validity of the list, we first duplicate the findings using the same data and calculations. Furthermore, since the data from the DMDC comes directly from DFAS, it is considered accurate for analysis.

The NSC database contains case information including detailed information about the situation in which the employee sustained the injury. It also includes demographic data such as age, gender and job type.

1. DMDC

In 1974 the DMDC was established as a Manpower Research and Data Analysis Center. DMDC was originally a DoD activity within the USN. Through the years it went under several changes. In 1997 DMDC was made part of a DoD Field Activity called the Defense Human Resources Activity (DHRA). According to the DMDC profile website, the DMDC's primary function has always been to support the information management needs of the Office of the Under Secretary of Defense for Personnel & Readiness (OUSD/P&R) and its predecessors. (DMDC, 2007) The mission of the DMDC is several fold:

---

[5] United States Department of Justice, <http://www.usdoj.gov/oip/privstat.htm>, 11 August, 2007.

- Collect and maintain an archive of automated manpower, personnel, training, and financial databases for the Department of Defense;
- Support the information requirements of the OUSD/P&R and other DoD manpower, personnel, and training communities with accurate, timely, and consistent data; and
- Operate DoD-wide personnel programs and conduct research and analysis as directed by the OUSD/P&R.
- To serve as the current repository for current and historic DoD human resource information
- To collect, provide and use this information to benefit decision makers in DoD and other government agencies, and members of DoD. (https://www.dmdc.osd.mil/pprofile/owa/intro.intro)

In order to receive data from the DMDC, an online account had to be established via the DMDC website. (DMDC, 2007) Once the online account was approved, an online request was made for the data. Despite the online request, several appointments were made with various DMDC personnel to solidify the details of the data request, but it was still time consuming. Reasons for the lengthy data request are mostly attributed to incomplete data extraction and difficulty matching employee LWD records and demographics when using scrambled identification numbers. Six different data sets provided by DMDC were found to be defective.

It is important to recognize that to study LWD rates one requires records of those employees who have at least one LWD and those employees who have no LWD. Therefore all requests for data must include records for all employees regardless of the number of LWD. Extracting these data proved difficult and time consuming for DMDC.

The data from the DMDC was received as a semicolon delimited text file on encrypted digital video discs. The LWD data was separated into two CY(s) per disc. Demographical data for each employee was captured once per CY for the month of May. The reason for this is file size and ease in manipulation records. Demographics for every USMC and USN employee would produce databases with eight gigabytes of data for each CY—too large for regular analysis. In addition, only choosing demographic records from May gives a unique set of demographic values per individual. The data was imported into Microsoft Access© where the fields were formatted for ease of use. Although Microsoft Excel© and S-Plus work well for data calculations and exploratory analysis, Microsoft Access© makes it easier to draw queries from the database which has 32 fields(columns) of data. The Dataset Management section of this chapter explains this process in detail. The fields used for analysis will be explained in the Exploratory Analysis section of Chapter 4. The field definitions can be found in Appendix A.

2. NSC

The NSC deals specifically with Naval and USMC safety and mishap data. Their mission is to provide safety assistance and advice to the Chief of Naval Operations (CNO), the Commandant of the Marine Corps (CMC), and the Deputy Assistant SECNAV for Safety in order to enhance the warfighting capability of the USN and USMC, preserve resources and improve combat readiness by preventing mishaps and saving lives. (NSC, 2007)

The functions of the NSC include guidance and direction, safety data services, safety program services, and marketing of safety. These functions support the overall Naval Safety Policy: to enhance operational readiness and mission accomplishment by establishing an aggressive occupational safety and health program. (Integrated Publishing, 2007). The Naval Safety Program is designed to reduce occupational injuries, illnesses, or death that could be brought upon by unsafe working conditions for all military service members and civilian employees.

Data from the NSC was received in two ways. The initial data was retrieved via the Web Enabled Safety System (WESS). The WESS provides an on-line, interactive, electronic means of managing mishap information and consolidating all types of incidents into one consolidated database at the NSC.

(NSC, 2007)[6] In order to receive data from the WESS, an account had to be authorized. Once this was completed, low level queries could be performed to retrieve the requested data from the NSC database. The files were saved in a Microsoft Excel column separated value (CSV) format.

Although the WESS allows for direct interaction with the NSC database, some of the more detailed fields are not made public to general WESS users. To alleviate this problem, the NSC was contacted. A database administrator (DBA) delivered pipe delimited text files via electronic mail (email).

The USMC NSC database consists of personnel tables and mishap log tables. These text files were imported into Microsoft Excel where the associated column/field names were assigned. The personnel and mishap log tables for the USMC consist of 79 and 97 fields respectively. These fields are defined in Appendix B.

The USN data was retrieved much in the same manner as the USMC data; however, there are differences in the way the NSC categorizes USN data. For example, instead of having one master database, the USN data was received as two separate databases: sea shore (SS) and motor vehicle (MV). Each database contains several tables which also contains several fields. The tables are: environment, reportable event, personnel, personnel injury, and type accident.

3. Dataset Management

In order to manage and use the databases from the NSC and DMDC, they were imported into Microsoft Access. All databases required additional configuration. Once the databases (one for the DMDC, three for the NSC) were imported into Microsoft Access, the tables for all associated databases had to be linked. For example, to link the tables for the USMC database, a field named "PersonTwoSN" (a unique identifier for each employee) was assigned as a

[6] Naval Safety Center, <http://www.safetycenter.navy.mil/wess/tutorial/default.htm>, 11 August 2007.

primary key within the personnel table. The primary key in the mishap log table, "mishap serial number", was inserted into the into the personnel table and marked as an alternate key. This alternate key acts as a "pointer" inside the mishap log table, consequently allowing a "one-to-many" relationship between the personnel table and the mishap log table. This means that one employee in the personnel table may have many mishap incidents within the mishap log table. The primary and alternate keys provided a means of connecting the two tables, enabling them to act as one database. This method was used on the two naval databases to achieve the same effect.

The DMDC database is different in structure than the NSC databases. There are two tables for each CY; a LWD table and a demographic table. The LWD table contains data for LWD computation and rates such as "Scheduled Days", "COP Days", "Pay Period", and "Base Name". Even though "Base Name" can be considered as a demographic for an employee, it is placed in the LWD table because LWD are not only assigned to a specific employee, but a specific base as well. This makes sense because an employee can be assigned to different bases during different pay periods. The structure of this table allows for direct LWD calculations for a specific base, employee, or pay period.

The DMDC demographics table contain fields such as "date of birth", "occupation code", "pay plan", and "pay grade". To connect the two tables, a unique identifier per employee was used. Therefore queries could be performed by matching the unique identifier in both tables.

4. Data Validation and Potential Problems

Data from the NSC and the DMDC have inconsistencies. The manner in which the NSC identifies employees is different than DMDC. The DMDC uses unique identifies for all employees, whereas the NSC uses case numbers to identify records. Employees are affiliated with case records. For the NSC, there can be more than one employee tied to a specific case. Without each employee having a unique identifier that matches DMDC identifiers, it is difficult to match employees.

The datasets from the DMDC and NSC are not consistent when considering the total LWD for the USMC per year. For example, a count of total LWD was conducted with the DMDC data and compared to the count of total LWD from the NSC dataset. For all years (CY 2000 – CY 2006) the total number of LWD are different. Table 3.1 shows the total number of LWD per year for the DMDC (all services) and Table 3.2 shows the total number of LWD per year for the NSC (USN and USMC).

Table 3.1 DMDC: Total number of LWD by CY.

	2000	2001	2002	2003	2004	2005	2006
USMC	3047.39	8396.08	10949.42	8651.76	7952.27	7939.82	5653.26
USN	22353.02	54309.27	69735.84	64320.76	58104.55	48018.85	41965.19

Table 3.2 NSC: Total number of LWD by CY.

	2000	2001	2002	2003	2004	2005	2006
USMC	2224	2285	3586	1729	803	2473	1560
USN	1191	662	470	1649	13585	17312	20633

Clearly the datasets contain different values of data for LWD. The data from the DMDC dataset show several thousands more LWD per year than the NSC dataset. There is no definitive answer to why, but it is assumed that the differences may depend on the reporting practices. For the NSC dataset, Terrie Rockett, the DBA responsible for providing the NSC data, stated that before FY 2003, the NSC only required reporting for incidents involving more than five LWD. Even after FY 2003, the number of LWD are still greatly different.

Another discrepancy is in the number of employees who have accumulated LWD. Tables 3.3 and 3.4 show the number of employees who have accumulated at least one LWD in each CY for the DMDC and NSC respectively. For example, CY 2005 shows the DMDC with 205 fewer employees who accumulated a LWD, while for CY 2003 the DMDC show 566 fewer employees who accumulated a LWD. This is after the aforementioned reporting procedure change. This is another indicator that the databases contain disparate data values.

Table 3.3 DMDC: Employees who have LWD by CY

	2000	2001	2002	2003	2004	2005	2006
USMC	40	169	427	345	283	304	273
USN	294	1054	2356	2333	2004	1693	1523

Table 3.4 NSC: Employees who have LWD by CY

	2000	2001	2002	2003	2004	2005	2006
USMC	277	230	375	1034	923	554	449
USN	95	87	67	94	1348	2378	1827

Another find was missing data values. Some records in both datasets had blank fields. Blank fields in the NSC database included: age, job title, pay grade, pay plan, unit code, and LWD. Blank fields in the DMDC database included: COP Date and Base Name. Those records with blank base names are not used.

A major point of concern is the interesting way in which the DMDC captures USMC employees. According to sources at the DMDC, initially employees are given an Agency Code. In the DMDC database, USMC employees fall under the DoN, therefore the code “NV” is applied as an Agency Code for all USN and USMC employees. According to Reza Nouri of the DMDC, in order to differentiate USMC employees from Naval employees, the DMDC uses a Major Command Code (MCC). The DMDC defines MCC as a code that designates the agency sub element to which an activity is assigned. (Appendix A) The MCC used to differentiate USMC employees is “27”. The problem with this code is that it is not an MCC used by the USMC. The major questions are,

1) does the “NV” Agency Code pose a problem by incorrectly identifying USN employees as USMC employees and vise versa, and 2) where does the MCC “27” get injected into the DMDC’s dataset?

A small comparison was done between Camp Pendleton, Camp Lejuene, MCLB Albany, MCLB Barstow, and MCLB Blount Island. Table 3.5 and Table 3.6 show the average age and standard deviation of employees who have a LWD. The purpose of this comparison is to show that not only is the LWD data for computation of the rate different between the datasets, but demographical data for these same employees are different as well.

Table 3.5 Comparison across USMC Commands DMDC (CY2006)

	Camp Pendleton	Camp Lejuene	MCLB Albany	MCLB Barstow
Avg Age/ Std Dev	49.42 / 9.79	49.18 / 10.12	48.21 / 9.34	48.65 / 9.98

Table 3.6. Comparison across USMC Commands NSC (FY2006)

	Camp Pendleton	Camp Lejuene	MCLB Albany	MCLB Barstow
Avg Age/ Std Dev	45.57 / 9.86	53.66 / 5.51	44.94 / 11.5	46.2 / 13.75

There is a problem with assigning civilian employees to bases. Initially, the data for some were combined due to proximity. For example, initially the data for USMC Air Station (MCAS) Beaufort and Marine Corps Recruiting Depot (MCRD) Parris Island were combined. The bases are about 10 miles apart and bring up questions of data integrity. According to Debbie Eitelburg, an employee of the DMDC, LWD are assigned to a specific base by zip code and geographic region. LWD accrued by employees are assigned by the employees’ current location (zip code), rather than unit address in which he is assigned. These zip codes are mapped to a specific base and therefore drive how LWD are assigned. While attempting to validate this claim, a query was performed to find out if MCAS Beaufort and MCRD Parris Island shared employees with the same zip codes. Employees with zip codes 29401 and 29901 showed up on both bases. Since employees with these zip codes did not contribute to the LWD count and these are two different types of bases, the bases were not combined.

Due to the discrepancies and differences with both databases, only the DMDC database was used for most of the analysis in Chapter IV. CY 2006 is the most recent year and thus was the focal point of the analysis.

IV. DATA ANALYSIS

The data analysis starts with exploration of the variables used in this study. Next, exploratory analysis is given followed by a chi-squared test for homogeneity that compares the USN and USMC employee distributions. A CT is fit to identify which groups of USMC employees are most likely to have at least one LWD. It is verified by a generalized additive model (GAM) and compared to a USN CT. The chapter concludes with a summary of the analysis and results.

A. EXPLORATORY ANALYSIS

1. Variables

a. Individual LWD Rate

To compare individuals it is convenient to use the LWD rate for each individual employee. The Individual LWD (ILWD) rate is the ratio of LWD to prorated days for each individual employee. Hence, for a particular year, the i^{th} individuals' ILWD is computed as:

$$ILWD_{rate} = \frac{\sum_j (COP + LWOP)_j}{\sum_j (\text{prorated days})_j} \tag{4.1}$$

where j indexes the pay period.

b. Independent Variables

The independent variables for this study include the numeric data values age (Age) and pay grade (PayGrade). The categorical independent variables for this study are base name (BaseName) and civilian occupation code (COC). These variables are found in the DMDC database. Pay Grade is an indicator of pay position covered by the same pay system. BaseName is the installation where the employee worked during a specific pay period. COC is the code given to a specific job. In addition, two more variables are constructed for the purpose of this analysis; base type (BaseType) and job type (COCCode). They are logical groupings of BaseName and COC respectively. They are also defined later in this chapter.

2. Descriptive and Exploratory Statistics

a. Exploration of USMC Bases

Table 4.1 ranks the USMC bases by LWD_{rate} in descending order. There are 20 specific USMC bases on file for CY 2006 that who had civilian employees that accrued LWD. In addition, approximately 200 employees were not assigned to a base. Their LWD rate per hundred employees is 76.6. To analyze LWD effects on specific types of bases, each base was given a code called BaseType based on functionality. The BaseType codes are: “LOG” for logistics, “AIR” for air station, “GROUND” for standard base, “ADMIN” for bases that are primarily for administration and management, “RDEPOT” for recruiting depot, and “UNK” for bases classified as reserve or unknown.

Table 4.1 LWD_{rate} for USMC Bases (CY 2006)

BaseName	LWD_{rate}	BaseType	LWD	Approx. #Emp*
BARSTOW MCLB	124.5514	LOG	1916.1	1537
NEW LONDON NAVSUBBASE	64.3399	OTHER	29	45
MCBH KANEOHE BAY	56.7291	GROUND	257	452
MAKIM,OKINAWA CP BUTLER	49.6210	GROUND	260	524
BEAUFORT MCAS	47.4675	AIR	140	295
PARRIS ISLAND MCRD	46.5409	RDEPOT	202	434
CAMP PENDLETON	46.2984	GROUND	761	1643
CHERRY POINT MCAS	30.8160	AIR	288	934
YUMA MCAS	29.2686	AIR	123	420
MCCDC QUANTICO VA	28.5708	ADMIN	380	1329
CAMP LEJEUNE MCB	26.0706	GROUND	434	1663
29 PALMS MC A/G CMBT CTR	22.4502	GROUND	163	725
MCSA KANSAS CITY MO	21.1336	ADMIN	41	194
ALBANY MCLB	17.5654	LOG	334.13	1902
HQTRS MARCORPS	7.7543	ADMIN	145	1870
CAMP H. M. SMITH	6.4724	GROUND	5	77
NEW ORLEANS NAS JRB	4.7436	OTHER	4	84
SAN DIEGO MC RD	1.9225	RDEPOT	7	364
IWAKUNI MCAS	1.5567	AIR	2	128
NEW RIVER MCAS	1.2327	AIR	2	162

*Note: The approximate number of employees given is an approximation based on total scheduled hours for that base divided by 260.

The most interesting characteristic Table 4.1 displays is that the LWD_{rate} do not appear to be driven by the type of base. The initial assumption was all LOG and GROUND type bases would yield the highest rates.

Outliers are data values that fall well outside the overall pattern of the data. (Weiss, 1995) There are five employees, each with 200 or more LWD for CY 2006, that are considered outliers. Upon examining these employees, four accumulated 130 or more LWD in CY 2005, therefore at various portions of analysis these employees are removed from the data set. It is assumed that these four employees had illnesses that carried over from CY 2005 and should not be included within the analysis for CY 2006. One employee had only 40 LWD in CY 2005; however, this employee accumulated the maximum amount of LWD amount starting in November. This pattern continued throughout CY 2006. One employee accumulated 396 LWD. During a regular CY a regular employee can only accrue 260 LWD [26 pay periods * 10 days per pay period]. An annual LWD total greater than 260 would typically be considered a data entry error or computational error; however, this is not the case. This employee has a COC of 00081 (Fire Protection and Prevention). The NSC made the following statement in a report on MCLB Barstow:

> Because fire fighters work a 24 hr shift, when one is injured, DOL charges three days for every shift a fire fighter is off work. For example, in 2006, Barstow's 129.5 lost work days due to fire fighter injuries equated to 425.5 in [the] DMDC [database]. In order to normalize the DMDC rate, the hours used in the calculation [RegHours] would have to be tripled to reflect the 24 hour shift. (NSC, 2004)

Upon further investigation of Fire Protection and Prevention employees, it was found that the DMDC uses these prorated hours in computing their LWD rates. It is important to know what any base where a Fire Protection and Prevention employee will have an inflated LWD total.

The significance of these five employees is that they contribute 44.01% of all LWD accumulated by USMC civilian employees for CY 2006. Table 10 shows LWD_{rate} excluding these five employees.

Table 4.2 LWD_{rate} for USMC Bases (CY 2006) excluding Outliers

BaseName	LWD_{rate}	Type	LWD	Approx. #Emp
BARSTOW MCLB	94.7597	LOG	1916.1	1537
NEW LONDON NAVSUBBASE	64.3400	ADMIN	29	45
MCBH KANEOHE BAY	56.7292	GROUND	257	452
BEAUFORT MCAS	47.4675	AIR	140	295
PARRIS ISLAND MCRD	46.5410	RDEPOT	202	434
CHERRY POINT MCAS	30.8161	AIR	288	934
YUMA MCAS	29.2687	AIR	123	420
MCCDC QUANTICO VA	28.5708	ADMIN	380	1329
CAMP LEJEUNE MCB	26.0706	GROUND	434	1663
29 PALMS MC A/G CMBT CTR	22.4502	GROUND	163	725
MCSA KANSAS CITY MO	21.1336	ADMIN	41	194
ALBANY MCLB	17.5654	LOG	334.13	1902
CAMP PENDLETON	10.0525	GROUND	761	1643
HQTRS MARCORPS	7.7543	ADMIN	145	1870
CAMP H. M. SMITH	6.4725	GROUND	5	77
NEW ORLEANS NAS JRB	4.7437	ADMIN	4	84
SAN DIEGO MC RECRUIT DEPOT	1.9226	RDEPOT	7	364
IWAKUNI MCAS	1.5567	AIR	2	128
NEW RIVER MCAS	1.2328	AIR	2	162
MAKIM,OKINAWA CP BUTLER	0.0000	GROUND	260	524

Highlighted in Table 4.2 are three bases affected by outliers; MCLB Barstow, Camp Pendleton, and Camp Butler, Okinawa. Once the outliers are removed, the LWD_{rate} decreases. MCLB Barstow’s rate dropped by 29.79 (remained in first position). Camp Pendleton’s rate decreased by 36.25 (dropped from position 8 to position 13). Camp Butler’s rate dropped completely (drop from position 5 to last position). One employee at Camp Butler, Okinawa accounted for all accrue LWD. The significance of this table is to show for some bases, outliers affect LWD_{rate} significantly.

b. Exploration of COC

Table 4.3 gives the LWD_{rate} for the top 20 COC ranked in descending order by total number of LWD. As suspected, the mechanic and fire fighter/police oriented COC make up most of the LWD in CY 2006. An interesting find in the table is how the rate per hundred employees is affected by a small number of employees. For example, out of the top twenty COC, "Coal Handling" has the highest rate per hundred employees (1117.188). This number can be misleading as the "Coal Handling" COC has 4 employees with 44 total LWD between them for an average of 11 LWD per employee for CY 2006. The COC "Ordnance Equipment Mechanic" suffers the same effect.

Table 4.3 LWD_{rate} for COC(USMC)

JobTitle	LWD_{rate}	Count	LWD	se
Coal Handling	1117.2	4	44	1170
Ordnance Equipment Mechanic	1048	4	42	1040
Electronic Measurement Eq Mechanic	740.7	27	200	706
Recreation Specialist	262.3	102	268	252
Police	232.1	37	87	135
Production Control	183.6	65	120	181
Sandblasting	130.9	49	64.1	122
Electronics Mechanic	122.6	167	205	96
Motor Vehicle Operator	109.4	216	233	76
Fire Protection and Prevention	107	677	1264	36
Heavy Mobile Equipment Mechanic	105.6	832	880	36
Quality Assurance	97.6	74	72	61
Laboring	89.7	88	79	55
General Business and Industry	75.6	185	140	69
Maintenance Mechanic	71.8	336	242	24
Miscellaneous Clerk and Assistant	48.6	507	239	47
Human Resources Assistance	29	392	110	18
Information Technology Management	24.5	817	200	27
Painting	24.4	172	42	22
Supply Clerical and Technician	16.3	276	45	11

An estimate of the standard error is given as an indicator of plausibility of the LWD_{rate} calculations. The equation used for Table 4.4 is:

$$se = \frac{\hat{\sigma}}{\sqrt{n}} \qquad (4.2)$$

where $\hat{\sigma}$ is the sample standard deviation of the COC group and n is the number of employees in the group.

A more appropriate approximate standard error of the COC LWD_{rate} is an approximation equation given by Rice (1988). Note that for each COC the LWD_{rate} per employee is the ratio:

$$\mathrm{R} = \frac{\frac{\sum(\mathrm{COP} + \mathrm{LWOP})_i}{n}}{\frac{\sum(\text{prorated days})_i}{n}} \qquad (4.3)$$

where i represents the i^{th} individual within a specific COC. Both, the numerator and the denominator in Equation (4.3) are statistics and therefore the standard error Equation (4.2) derived for a sample mean is not necessarily appropriate. Instead, Rice (1988) gives an approximate standard error for a ratio:

$$\mathrm{R} = \frac{\frac{\sum \mathrm{Y}_i}{n}}{\frac{\sum \mathrm{X}_i}{n}} \qquad (4.4)$$

where $Y_1, Y_2, \ldots, Y_n$ and $X_1, X_2, \ldots, X_n$ are random variables representing LWD and prorated days for the n individuals in a specific COC. Let μ_Y, μ_X, σ_Y, σ_X be the expected values and standard deviations and σ_{XY} be the covariance of X and Y. Furthermore, suppose $Y_1, Y_2, \ldots, Y_n$ are independent and $X_1, X_2, \ldots, X_n$ are independent, then (ignoring the finite population correction factor):

$$Var(R) \approx \frac{1}{n{\mu_X}^2}\left(r^2\sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_{XY}\right) \qquad (4.5)$$

where the population correlation coefficient ρ is defined as:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{4.6}$$

and r is defined as:

$$r = \frac{\mu_Y}{\mu_X} \tag{4.7}$$

To get the subsequent standard error for the ratio, replace the population parameters with sample estimates and multiply the standard error by 26000. The "26,000" scaling factor represents hundred employees working 10 days for 26 pay periods. Note that for the rates in Table 4.4, the standard errors computed in Equations (4.2) and (4.5) are almost equal. In these cases, the more simple approximation Equation (4.2) is used. In general, COC with smaller sample sizes, for example Coal Handling from Table 4.4, tend to have larger standard errors.

MCLB Barstow has an adjusted LWD_{rate} of 94.8 (adjusted for outliers) compared to MCLB Albany with a LWD_{rate} of 17.6. Both large logistics bases are often compared. There are a total of 181 COC between both bases. The two have 109 COC in common (approximately 60%). Albany has 46 COC that Barstow does not have. Barstow has 26 COC that Albany does not have. This is a good indicator that the distribution of employees at the bases is somewhat similar, but not identical. However, this observation changes when considering which COC contribute to the associated LWD_{rate}. Table 4.4 shows an analysis of COC that incurred LWD for MCLB Barstow and Albany.

Table 4.4 Differences in COC for MCLB Albany and MCLB Barstow

	Albany			Barstow		
JobTitle	Count	LWD	LWD_{rate}	Count	LWD	LWD_{rate}
Heavy Mobile Equipment Mechanic*	353	123	34.84	412	441.1	113.30
Maintenance Mechanic	11	2	18.36	23	92	392.65
Painting	81	3	3.70	51	36	70.03
Welding	18	1	5.55	24	1	4.23
Electrical Equipment Repairer	51	7	13.70			
Electrician	5	2	39.33			
Electronic Integrated Systems Mech	15	5	33.90			
Electroplating	6	31	519.66			
Inventory Management	75	2	2.68			
Machining	32	4	12.55			
Miscellaneous Clerk and Assistant	16	80	506.70			
Misc Electrical Installation and Maint	4	1	24.98			
Misc Industrial Equipment Operation	1	2	198.47			
Sandblasting	38	64.1	168.73			
Small-Arms Repairing	13	1	7.83			
Supply Program Management	27	6	22.30			
Air Conditioning Equipment Mechanic				5	30	600.00
Artillery Repairing				29	11	38.38
Contracting				4	10	250.00
Education and Training Technician				10	2	20.70
Electronic Measurement Eq Mech*				16	0	0.00
Electronics Mechanic				62	202	326.58
Engineering Technician				17	3	17.72
Environmental Protection Specialist				24	1	4.19
Equipment Cleaning				19	11	58.00
Fire Protection and Prevention				52	113	125.25
Fork Lift Operating				13	15	118.00
General Equipment Mechanic				13	15	115.56
Laboring				9	2	21.65
Library Technician				4	1	27.93
Locksmithing				1	5	634.15
Materials Handler				27	10	37.25
Mobile Equipment Metal Mechanic				21	3	14.33
Motor Vehicle Operator				12	160	1368.42
Ordnance Equipment Mechanic				4	42	1047.98
Packing				2	1	64.84
Police				21	86	411.33
Production Control				27	120	446.99
Quality Assurance				14	40	279.57
Railroad Repairing				3	2	66.50

As the table indicates, there are only four COC that can be directly compared. This is an indication that caution needs to be taken when comparing two bases that appear similar on the surface. Of those that accrued LWD, there are 36 out of 40 COC that MCLB Barstow and MCLB Albany do not have in common.

Since bases do not have the same COC and because the standard errors of LWD rates for COC with a small number of employees are high, a new variable COCCode is constructed. This variable combines COC into eight groups by general job type. Table 4.5 summarizes the job type for each COCCode. Appendix C gives the specific COCCode for each of the eight groups of Table 4.5

Table 4.5 COC Code

COC Code	General Description of Job Type
A	Fire Fighter/Police/Security Forces
B	Mechanic
C	Management/Administration/Finance/Business
D	Medical Fields
E	Equipment Operators
F	Information Technology/Science/Engineers
G	Aircraft
H	Miscellaneous

Table 4.6 shows a comparison of the number of employees, LWD_{rate}, number of LWD, and average age with the associated standard deviation for each COCCode.

Table 4.6 Differences in COC for MCLB Albany and MCLB Barstow

	Albany								Barstow						
COC Code	% Emp	# Emp	LWD_{rate}	LWD	% LWD	Avg Age	StDev		% Emp	# Emp	LWD_{rate}	LWD	% LWD	Avg Age	StDev
A	1%	17	0.0	0	0%	43.6	8.0		5%	73	178.6	199.0	10%	38.1	10.5
B	26%	491	26.5	130	39%	47.1	11.2		42%	608	211.5	1287.1	67%	42.9	13.2
C	30%	559	15.7	88	26%	51.2	7.7		13%	184	92.7	170.0	9%	50.2	8.7
D	0%	9	0.0	0	0%	51.8	7.8		1%	7	0.0	0.0	0%	43.0	5.5
E	8%	154	46.2	71.1	21%	49.4	10.9		7%	108	163.1	176.0	9%	47.4	10.0
F	13%	240	4.2	10	3%	48.4	9.7		9%	128	2.3	3.0	0.2%	48.2	11.9
H	21%	389	9.0	35	10%	48.3	10.4		23%	338	23.7	80.0	4%	47.9	10.9

Table 4.7 shows the results in Table 4.6 excluding outliers.

Table 4.7 Differences in COC for MCLB Albany and MCLB Barstow minus Outliers

	Albany								Barstow						
Code	% Emp	# Emp	LWD_{rate}	LWD	% LWD	Avg Age	StDev		% Emp	# Emp	LWD_{rate}	LWD	% LWD	Avg Age	StDev
A	1%	17	0.0	0	0%	43.6	8.0		5%	73	178.6	199.0	14%	38.1	10.5
B	26%	491	26.5	130	39%	47.1	11.2		42%	606	136	827.1	57%	42.9	13.2
C	30%	559	15.7	88	26%	51.2	7.7		13%	184	92.7	170.0	12%	50.2	8.7
D	0%	9	0.0	0	0%	51.8	7.8		1%	7	0.0	0.0	0%	43.0	5.5
E	8%	154	46.2	71.1	21%	49.4	10.9		7%	108	163.1	176.0	12%	47.4	10.0
F	13%	240	4.2	10	3%	48.4	9.7		9%	128	2.3	3.0	0.2%	48.2	11.9
H	21%	389	9.0	35	10%	48.3	10.4		23%	338	23.7	80.0	5%	47.9	10.9

Clearly in almost every group, MCLB Albany and MCLB Barstow are numerically different. Note in five of the seven groups MCLB Albany has older employees. Also note the difference in Code B (Mechanic Group) employees. Group B for Barstow makes up 42% of its workforce and only 26% for Albany. However, when considering the percent of LWD, they are comparable. The largest differences appear to be in Codes A and C. Code A makes up 27% more LWD for Barstow whereas Code C makes up 12% more for Albany. While the missions and total COC composition of the bases are similar by approximately 60%, the composition of employees COC that contribute to the LWD_{rate} are completely different.

c. *Comparing USMC and USN*

Next, LWD_{rate} between the USMC and USN were compared. The DMDC has records for approximately 17,470 and 170,251 civilian employees for the USMC and USN respectively. In CY 2006 the USN LWD_{rate} was 26.1 while the USMC LWD_{rate} was 35.2. To study these rates more carefully the distribution of individual LWD is compared. Since MS Excel cannot perform analysis on datasets with over 65,569 records, a query was performed on USN data and subsequently transferred into S-Plus. With S-Plus, a random sample of 17,470 Naval records was selected and transferred to MS Excel for analysis.

Table 4.8 shows the frequency and cumulative distribution of individual LWD for USMC and USN civilian employees. The table measures the frequency or number of employees whose $ILWD_{rate}$ falls within the indicated bin.

Table 4.8 Tabular Histogram USMC vs. USN

USMC			USN		
$ILWD_{rate}$	*Frequency*	*Cumulative %*	*$ILWD_{rate}$*	*Frequency*	*Cumulative %*
0	17197	98.44%	0	17323	99.16%
0-0.08	215	99.67%	0-0.08	99	99.73%
0.081-0.16	29	99.83%	0.081-0.16	22	99.85%
0.161-0.24	10	99.89%	0.161-0.24	6	99.89%
0.241-0.32	6	99.93%	0.241-0.32	1	99.89%
0.321-0.4	1	99.93%	0.321-0.4	3	99.91%
0.41-0.48	2	99.94%	0.41-0.48	1	99.91%
0.481-0.56	2	99.95%	0.481-0.56	4	99.94%
0.561-0.64	1	99.96%	0.561-0.64	1	99.94%
0.641-0.72	1	99.97%	0.641-0.72	0	99.94%
0.721-0.8	1	99.97%	0.721-0.8	0	99.94%
0.81-0.88	0	99.97%	0.81-0.88	2	99.95%
0.881-0.96	0	99.97%	0.881-0.96	0	99.95%
0.961-1	5	100.00%	0.961-1	8	100.00%
Total	17470		Total	17470	

The largest disparities can be seen in employees whose $ILWD_{rate}$ are between 0 and 0.16. This means that compared to the USN, the USMC has more employees with a lower $ILWD_{rate}$. To support this claim a Chi-Squared test was performed to test the following null hypothesis:

H_o: $ILWD_{rate}$ distributions for the Navy and Marine Corps are the same

To perform this test, a table of estimated expected values, under the null hypothesis, is computed. Let the expected number of employees who fall into cell ij be denoted by E_{ij} where j indexes USMC ($j=1$) and USN ($j=2$) and i indexes the bin $i = 1, 2,\ldots,14$. Let $O_{.1}$ and $O_{.2}$ denote column totals for the USMC and USN respectively. Likewise, let $O_{1.}, O_{2.},\ldots,O_{I.}$ denote row or bin totals. Under the null hypothesis the probability that an employee for each column j (USN or USMC) fall into row i ($ILWD_{rate}$ bin) is the same. Therefore,

$$E_{ij} = \frac{\left(O_{i.} * O_{.j}\right)}{O_{..}} \tag{4.8}$$

According to Navidi (2006), the Chi-Squared test statistic is based on the differences between the observed and expected values and can be computed as:

$$\chi^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{4.9}$$

Under H_o, the distribution of this test statistic has an approximate Chi-Squared distribution with (I-1)(J-1) degrees of freedom. For analysis, bin 0.96 had to be removed because no employee, USMC or USN, fell into this bin; it causes division by zero in the Chi-Squared test statistic.

The results yielded a test statistic of 55.54 and with 12 degrees of freedom and p-value of $1.44e^{-7}$. Therefore at any reasonable level of significance, the null hypothesis is rejected. From the chi-squared test there is significant evidence that the distributions of ILWDrate are different between the USMC and USN.

Further, the greatest contributions to the Chi-Squared statistic are from the first two bins of the distribution. The contributions from bins with ILWD rate greater than 0.08 are negligible. Thus the greatest difference between USN and USMC are in the proportion of employees who exhibit at least one LWD.

The USMC has a higher proportion of employees (1.5%) who have at lease one LWD in CY 2006 than the USN (0.9%). However, the average ILWD rate among those with one LWD is only 6.7 per employee for the USMC and 9.6 per employee for the USN. It can also be seen in Figure 4.4 which gives the frequency distributions of ILWDrates for those individuals who have at least one LWD. This suggests that USN has a higher LWD rate among employees that have at least one LWD.

Based on these results, the analysis now focuses on the proportion of employees who have at least one LWD. This has two advantages: 1) it focuses attention on the real difference between the USN and the USMC and 2) it gives results that are more robust to a few individuals with extreme total LWD. This is particularly important when we start considering LWD by COC, age, base etc.

B. CLASSIFICATION TREE

1. Methodology

The second method of analysis used is a CT; a hierarchical display of classifications based on a series of questions asked about each unit in a sample. (Montgomery et al., 2000) The structure of the tree is based on the order in which the questions are asked. According to Montgomery et al., "the general principle is to ask the question that maximizes the gain in node purity at each node-splitting opportunity…" Node purity refers to the level of variability in the response variable at each node within the tree. For our purposes, the response variable is the binary variable which takes value "1" if an employee has at least one LWD in CY 2006 and "0" otherwise. For a binary response, the measure of impurity is "the measure of node purity or deviance at each node … just the corrected sum of squares of the observations at that node". (Montgomery et al., 2000)

The shape of the tree begins with a root node that contains all of the observations within the sample. From this node and each subsequent node, the CT algorithm answers two questions: “1) which variable to spit one and 2) at what threshold [deviance].” (Conatser, 2006) Note that these questions are asked recursively at each subsequent child node throughout the tree structure. The goal is to “produce the effect of partitioning the data at an internal node in two disjoint subsets (branches) in such a way that the class labels [classification] in each subset are as homogeneous as possible.” (Hand et al., 2001) Theoretically, a CT can model the entire dataset, classifying each observation of data according to it attributes. However, this would produce an over-fitted CT that will be difficult to interpret and have diminished predictive ability. Pruning helps avoid this situation.

“Prune” means to reduce the size of the tree by removing nodes that contribute the least in predicting. (Conatser, 2006) The CT uses a greedy heuristic algorithm to recursively “grow” a tree from the root node. Based on pruning methods, we to scale down the size of the tree to increase accuracy in identifying of predicting outcomes of unseen samples.

2. Analysis and Results

For the purpose of implementing the CT, S-Plus has a “tree” function that performs the recursive partitioning algorithm. Figure 4.1 plots a cross-validated estimate of deviance against the size or number of leaves of the tree. This plot indicates that eight may be the best size tree to use. Note that a tree of say, 20 or larger, does not yield a tree with lower deviance than a tree of size five.

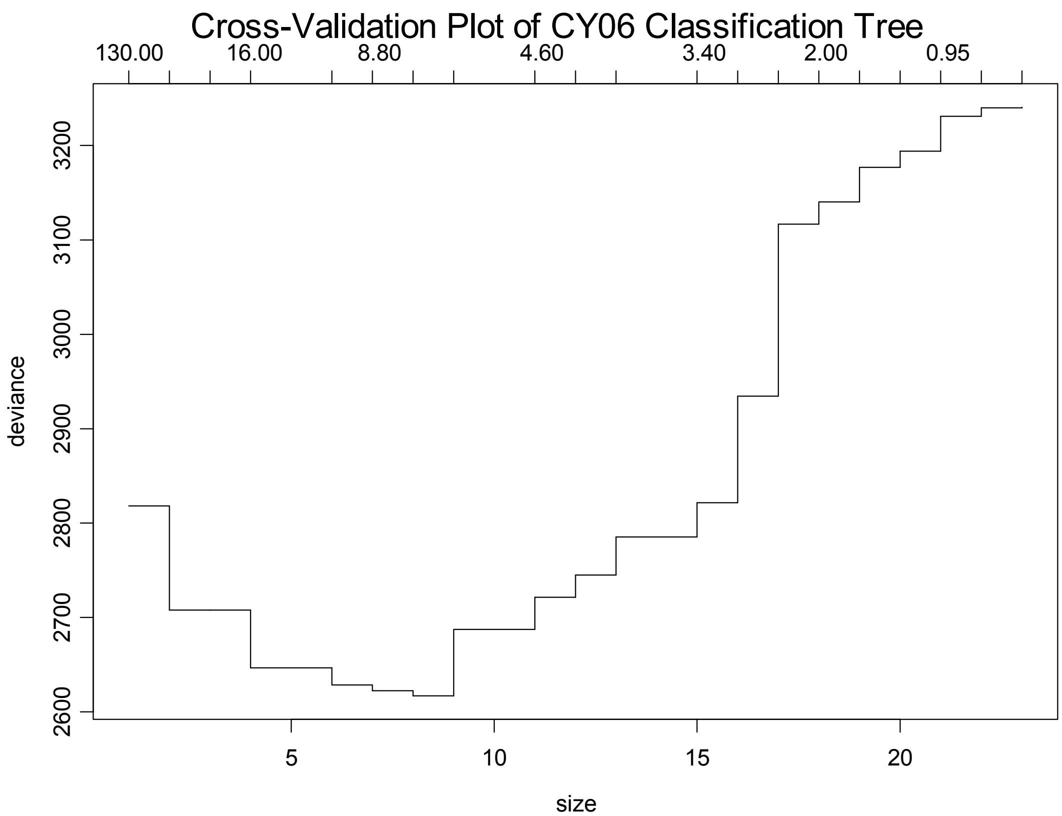


Figure 4.1 Cross-Validation Plot of CY06 CT Tree (USMC)

Figure 4.2 shows the resulting tree for all USMC employees during CY 2006. Using the variables described in Section A, we find the CT algorithm uses only COCCode, BaseType, and PayGrade to make predictions. Notice under the root node "273/17470"; this number means 273 out of 17470 employees actually accrued at least one LWD. The primary split occurs within COCType. Codes C, D, F, and G split to the left and the remaining codes to the right. The zero followed by the percentage within the nodes (circles) indicate that out of all the categories, the CT algorithm did not predict any employees to accrue a LWD. In this case, since the response variable is binary or dichotomous, a "1" would indicate that with a new dataset, it is predicted that an employee in that particular node is most likely at risk of accruing a LWD. However, the associated percentage gives the reader and indication of which node (category) an employee is most likely to accrue a LWD. The terminal node (rectangle) with 4.5% indicates that employees who are fire-fighters, security forces, and mechanics are at the greatest risk for accruing a LWD.

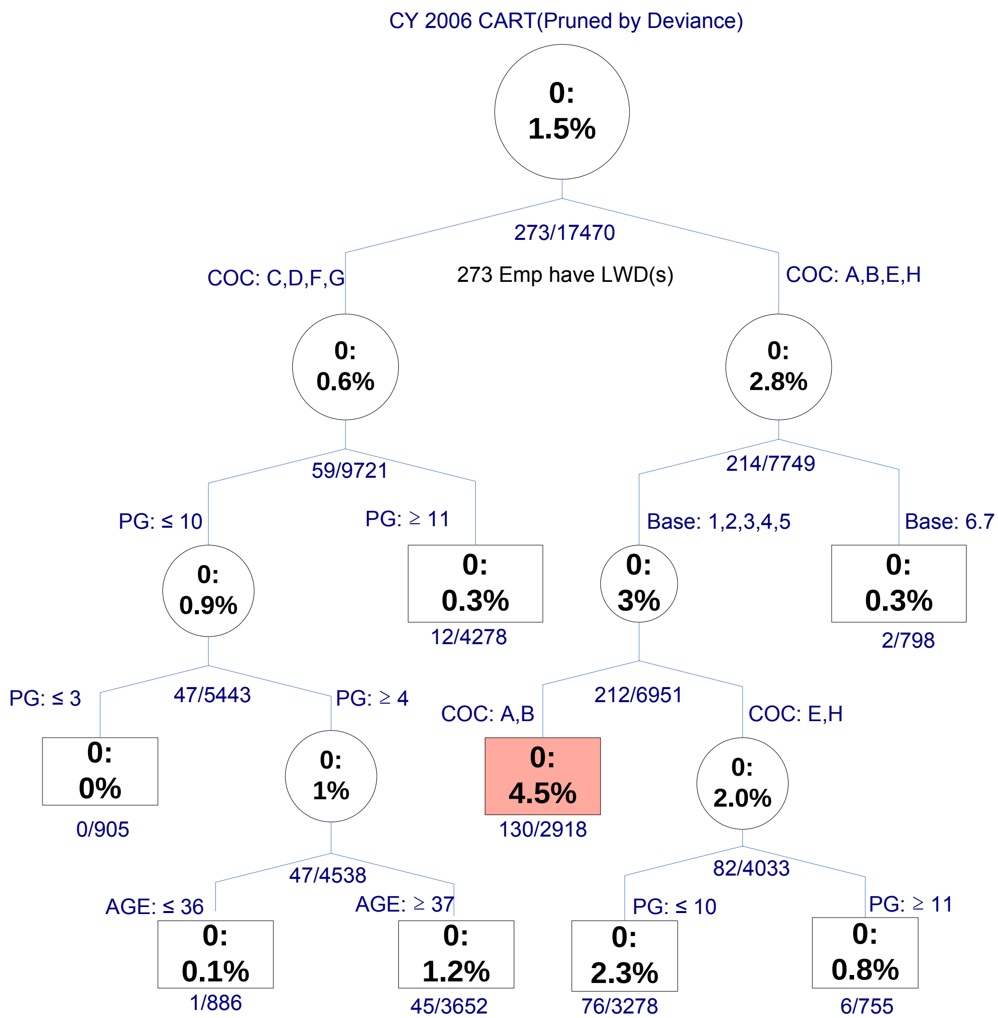


Figure 4.2 CY 2006 CART(Pruned by Deviance)

C. GENERALIZED ADDITIVE MODEL (GAM)

The largest difference between the USN and the USMC appears to be how many employees accrue a LWD. The CT shows what variables contribute to an employee accruing a LWD. To validate these findings a GAM is used. In this model, the binary response variables indicating employees who have at least one LWD are modeled as independent Bernoulli variables with probability of at least one LWD π_i for the i^{th} employee. The probability π_i is "linked" to the explanatory variables through a link function which in this case is the logistic link function $\log\left(\pi / 1-\pi\right)$. For exploratory purposes, the logistic link is additive in the independent variables with no interactions. Smoothing splines are fit to capture the non-linear contribution of the numeric variables "PayGrade" and "Age".

Appendix C gives the details of this GAM fit. Included are partial residual plots including the partial fits for each independentjm variable. They confirm the tree results, namely middle pay grades, COCCodes A, B, E, H, and higher ages contribute to a higher proportion of people with LWD.

D. OTHER ANALYSIS

A possible reason for the USMC not meeting the SECNAV's LWD_{rate} goal is the rise in employee numbers. Figure 4.3 shows the percentage increase or decrease in total LWD and employees based on the previous year (2002 – 2006).

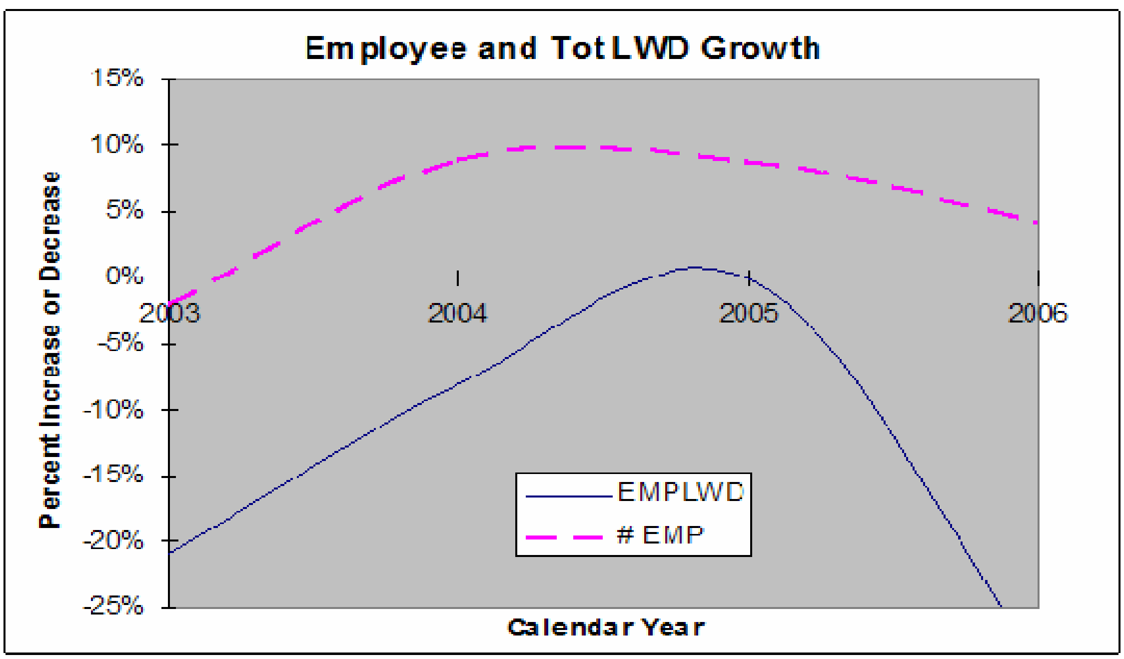


Figure 4.3 Percent Incr/Decr in Employee Number and LWD

Since the base year of 2002, the percent change in LWD totals have generally matched increases and decreases in the percent change in the total number of employees. The graph shows that an increase in employees can be directly related to an increase in LWD. For example, throughout early 2003, the percent change in number of employees increased. Likewise the percent change in LWD totals increased in this same period. From 2005 to 2006 the percent change in the number of employees decreased along with the percent change in the number of LWD. However, an increase in total LWD does not necessarily imply an increase in the LWD rate. From 2003 to 2006, the civilian workforce has risen from 14750 to 17845, a total of 17.3%. Despite this significant increase, the USMC has decreased the annual LWD_{rate} by 10%.

In Chapter IV, Section 2.b, a Chi-Squared test showed the individual-employee LWD rates between the services (USN and USMC) are different. By using a histogram, personnel LWD rates for the entire population of USMC employees were compared to those from an equally-sized random sample of USN employees. This comparison was based on $ILWD_{rate}$. The greatest disparities can be found in those who have $ILWD_{rate}$ less than 0.16. This group consists of approximately 244 USMC employees and 121 USN employees (Figure 4.4). The USMC has a larger percentage of employees that incur LWD; however, USN employees produce more LWD per individual.

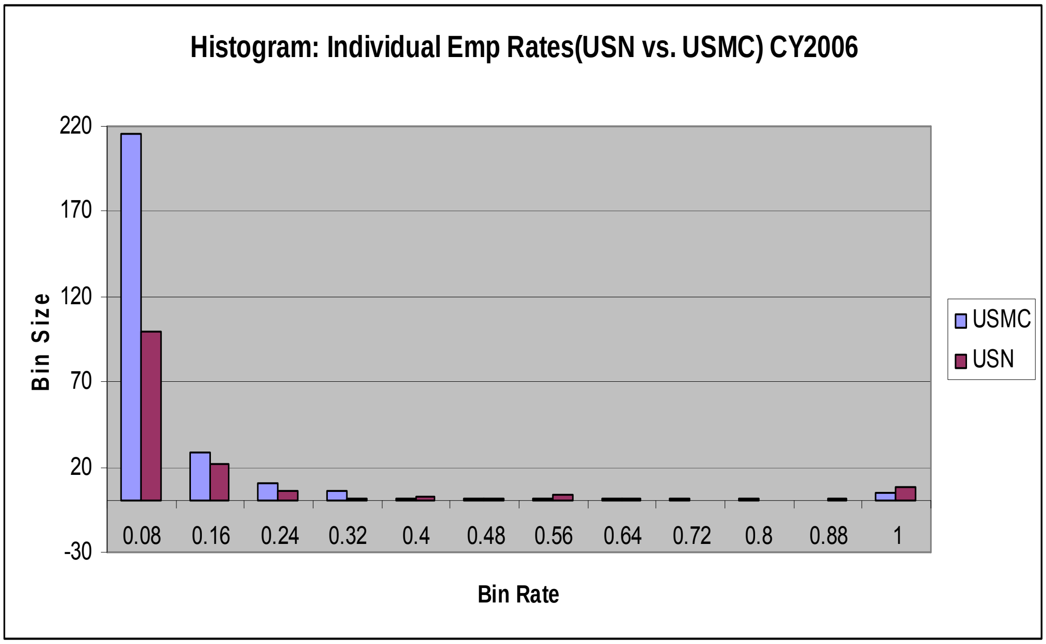


Figure 4.4 Histogram of $ILWD_{rate}$ USN vs. USMC

The CT of Figure 4.2 indicates that for the USMC the high risk COCCodes are A and B, the moderate risk COCCodes are E and H, and all others are low risk. In table 4.9, we compare these groups to like USN groups. Table 4.9 is constructed from all USMC civilian employees and all USN civilian employees with like COC as USMC civilians. The groups are derived from the major splits in the CT (Figure 4.2).

Table 4.9 High Risk Group Percentages USN vs. USMC

		Totals		**LWD Percentage**	
COC Type	Risk	USN	USMC	USN	USMC
A, B	High	10.5%	19.3%	2.7%	3.9%
E, H	Med	21.1%	25.6%	1.6%	1.8%
C, D, F, G	Low	68.4%	55.1%	0.45%	0.6%

There are more USMC civilian employees in the high risk groups and fewer in the lower risk groups. For example, for fire fighters, police, and mechanics, the USMC has 8.8% more employees than the USN in this category. Consequently, the percentage of USMC LWD in this category is higher than the percentage of USN LWD as well.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSION

This thesis is part of a continuing investigation into USMC civilian LWD. It was briefed two times to the Executive Safety Board, chaired by the Assistant Commandant of the Marine Corps—once at the start of this study and then again at the conclusion.

The USMC has not achieved the SECNAV's LWD rate goal. A possible reason for this is the rise in employee hiring. Since the base year of 2002, the percent change in LWD totals have increased and decreased with respect to the percent change in the total number of employees.

A notable find stated in Chapter 3 is 44% of all LWD recorded for CY 2006 were from employees who were accruing LWD in CY 2005. It is assumed that these LWD were from prior injuries. Therefore, two metrics would better represent the scale of LWD: adjusted LWD rate (ALR) and Focused LWD rate (FLR). ALR would focus all employees that accrued LWD during the current year, whereas FLR use a percentage of those who accrued at least one LWD. The LWD rate without outliers for CY 2006 is 26.76 per hundred employees; still higher than the 11.7 LWD_{rate} goal for FY 2006, but a significant 36% drop from FY 2002. In addition, those employees with LWD greater than 200 and LWD in the year prior should be considered on a case by case basis.

Of all employees who accrued a LWD, those who are fire fighters, security forces, and mechanics have the highest probability of accruing a LWD. The next group of concern is equipment operators and those designated miscellaneous whose pay grade is lower than 10. Refer to Appendix E for COC designations.

The ILWD rates and types of employees between the services are different. It is important to note that the USN has 180 more COC than the USMC. The greatest disparities can be found in those who have ILWD rates less than 0.16. For CY 2006, this group consists of approximately 244 USMC employees and 121 USN employees taken from an equal sized random sample of USN employees. The USMC has a larger percentage of employees that incur LWD (1.5% USMC, 0.9% USN); however, USN employees produce more LWD per individual (6.7 USMC, 9.6 USN).

There is no certain type of USMC base that is prone to a high LWD_{rate}. Results in listed in Table 4.2, sorted in descending order of LWD_{rate}, show an even spread of all types of USMC bases. When comparing bases, it is important to take into account the distribution of employee COC. For example MCLB Barstow and MCLB Albany have similar missions, but their COC distributions are completely different. Of those COC Codes that contributed to the LWD rates of both bases, there are only four that the two bases share. Therefore it is difficult to compare these bases directly. Not only does COC type affect the overall LWD_{rate} for a base, but the number of employees with particular COC affects the LWD_{rate} as well. For example, MCLB Barstow has a higher LWD_{rate} than MCLB Albany, and MCLB Barstow has 47% of its civilian employees in the high risk COC groups A and B while MCLB Albany has 27% of its employees in these groups.

B. FUTURE RESEARCH AREAS

Future research areas should continue the investigation begun in the thesis. Extensions of this thesis should include investigation of variables such as prior service, gender, race, education level, job term, etc. Research should also include a more careful grouping of COC codes so that other services can be compared to the USMC. In addition, it should be investigated whether base LWD reporting practices and DMDC subsequent manipulation of the LWD records affect LWD rates. Furthermore, the ability for the NSC to match its records with DMDC records should be investigated.

The factors that influence LWD per injury should be examined in more detail. For example, types of injuries and return-to-work programs can be included to discern if there are specific injuries associated with a particular base or COC and to find out if a particular return-to-work program is effective enough to implement throughout the USMC.

Lastly, since there were no standing operating procedures for the reporting and tracking of LWD, research could be performed to find the most efficient and accurate way of monitoring LWD for individual commands and bases. Systems such as ERP, DCPDS, and SLDCADA can be examined to determine if there is a better way of reporting LWD. Improving the user interface of these software applications to decrease the number of key type errors can be examined with a human factors-based approach.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. LIST OF FIELDS FOR DMDC DATA

Variable Name	Type	Description
Identifier	Numeric	Identifier is a nine digit numeric identifier based on a scrambled SSN.
Regular Hours	Zoned Decimal	Number of Regular Hours worked for the pay period.
Regular Day	Zoned Decimal	The Regular Hours divided by 8 to convert to Days.
Date of Birth	Numeric	YYYYMMDD
Annual Leave Balance	Numeric	Annual Leave Balance as of this pay period.
Admin Leave	Numeric	Administrative Leave taken this pay period.
Salary	Zoned Decimal	Basic Salary
Pay Basis	Character	BW - Bi-Weekly FB - Fee Basis PA - Per Annum PD - Per Diem PH - Per Hour PM - Per Month PW - Piecework SM - Semi-Monthly SY - School Year WC - Without Compensation
Locality Pay	Zoned Decimal	Annual Amount of additional pay received based upon one's duty location.
Pay Plan	Character	A plan prescribed by law or other authoritative source that governs the compensation paid an employee. See Appendix A.
Pay Grade	Character	Graduated scale of ranks within a specified pay plan. See Appendix B.
Pay Step	Numeric	Increments of basic pay authorized by a pay plan within a specific grade. See Appendix C.
Employee Status Code	Character	The current status of an employee, whether active or inactive, receiving severance pay, assigned to light duty or on an approved long term absence. See Appendix D.
Gross Pay	Zoned Decimal	Includes basic, premium and any other pay and allowances for this pay period.
Pay Period End Date	Numeric	YYYYMMDD
Net Pay	Zoned Decimal	The Employee's gross wages reduced by all deductions applicable for the pay period (e.g. taxes, allotments, retirement, Medicare, garnishments). The actual amount received by the employee for the pay period.
Agency	Character	AF - Department of Air Force AR - Department of Army DD - Department of Defense (except Air Force,

Variable Name	Type	Description
		Army, and Navy) NV - Department of Navy
Major Command Code	Character	A code that designates the agency sub element to which an activity is assigned. (Also known as bureau code for civilians).
Work Schedule	Character	A code designating the employee's work schedule. F - Full time G - Full time - Seasonal H - Full time - On call I - Intermittent J - Intermittent - Seasonal P - Part time Q - Part time - Seasonal R - Part time - On call S - Part time - Job Sharer T - Part time - Seasonal Job Sharer
Amount of OWCP Hours	Zoned Decimal	The amount of hours being paid by the Department of Labor under Workers Compensation for the pay period.
Amount of OWCP in Days	Zoned Decimal	The OWCP Hours divided by 8 to determine days.
Amount of Regular Hours	Zoned Decimal	The amount of regular hours worked for the pay period.
Amount Regular Days	Zoned Decimal	The regular hours divided by 8 to determine days.
State of Residence	Character	The two character abbrivation for State that the US Postal Service uses.
Amount of Hours Lost Due to Injury	Zoned Decimal	Actual hours lost on the day of injury.
Amount Lost in Days	Zoned Decimal	The Hours Lost Due to Injury divided by 8 to determine days.
Retroactive Amount of Hours Lost Due to Injury	Zoned Decimal	For those that are reporting hours from a previous pay period.
Retroactive Amount Lost in Days	Zoned Decimal	Same as above divided by 8 to determine days.
Injury Date	Numeric	YYYYMMDD
Injury Number	Character	
COP Date	Numeric	YYYYMMDD, Date the Continuation of Pay began
COP Number	Character	Number Identifying COP case
Retroactive Injury Date	Numeric	YYYYMMDD, Date for those injuries that aren't reported in the same pay period that they happen.
Retroactive Injury Number	Character	Number identifying Retro Inj
Retroactive COP Date	Numeric	YYYYMMDD, Date of Continuation of Pay that wasn't reported in the same pay period that they happened.
Civilian	Character	Occupation code is submitted to DMDC by

Variable Name	Type	Description
Occupation Code		DoD agencies consistent with standards set by the U.S. Office of Personnel Management (OPM).
UIC	Character	The unique code that represents the unit organization that an active-duty sponsor is officially assigned to (AKA UIC).
Base Identification	Numeric	Numeric code which identifies a military installation.
Installation Name	Character	The Installation Name that applies to a specific Base Identification Code.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. LIST OF FIELDS FOR NSC DATA

PersonTwoSN	No definition given
MishapLogSN	number assigned to a mishap that is machine generated
FileNum	this field identifies the mishap number. Values: generally populated with RUC/MCC/date
InjClass	Injury category based on personnel's injury status
MCC	Monitor Command Code assigned to personnel injured or involved in mishap.
RUC	Reporting Unit Code assigned to personnel injured or involved in mishap.
CompCMD	Headquarters level command of personnel injured or involved in mishap. I.E. MARFORLANT, MARFORRES.
MajCMD	Major Command level of personnel injured or involved in mishap. I.E. II MEF, MC Bases Pacific
ParentCMD	Parent Command of personnel injured or involved in mishap. I.E. 2nd MAW, 3rd MARDIV
UnitCMD	Unit personnel injured or involved in mishap belonged to. III MEF HQ, MAG-29, 12th Marines, 8th COMM BN
CompDept	Unit, Squadron, Company or Department of Personnel injured or involved in mishap.
LastName	Last name of personnel injured or involved in the mishap
FirstName	First name of personnel injured or involved in the mishap.
MI	Middle initial of personnel injured or involved in mishap.
Rank	Two digit rank of personnel injured or involved in mishap. I.E. E-3, GS, O-5
MOS	MOS of personnel injured or involved in the mishap.
Rate	Rate of Navy personnel injured or involved in mishap. I.E. BMC, PRCM, etc.
Age	Age of Person Injured
Gender	Male or female status
BilletMOS	MOS at the time of the mishap – may not be permanent MOS
Personnel	Civilian
Status	Injury status of personnel injured. I.E. fatality, lost time, no lost time, first aid.
JobTitle	Job title of personnel injured or involved in the mishap
InjuryType	Type of injury personnel suffered at the time of the mishap, I.E. sprain/strain, abrasion, drowning

BodyPart	Major Body Part Injured at time of the mishap
Seatbelts	Checked if seatbelst were worn by personnel injured or involved in mishap
ReflectiveVest	Checked if reflective vest was worn by personnel injured or involved in mishap
FlakJacket	Checked if a flakjacket was worn
HelmetHardHat	Checked if head protection was worn a the time of the mishap.
Sleeves	No definition given
GogglesGlasses	Checked if eye protection was worn at the time of the mishap
Earplugs	Checked if hearing protection was worn
FlotationDevice	Checked if personal flotation device was worn
Pants	Checked if long legged pants were worn during the mishap
Gloves	Checked if safety gloves were worn at the time of the mishap
SafetyBoots	Checked if any safety boots were worn by personnel injured or involved in mishap.
OtherPPE	Typed in if other PPE not on list was worn
PPEIncorrectly	No definition given
Seated	Location in vehicle, equipment, or general area that personnel injured or involved in mishap was at the time of mishap. I.E. operator, passenger, pedestrian., bicyclist.
LWDStatus	No definition given
StatusCost	Cost of the injury based on computation from DOD 6055.7.
DateOfStatus	Date the person lost the first day of work or date of death. May be after the initial injury. I.E. date of death, date diagnosed as permanent partial disability, date the lost workday started.
OSHACodes	OSHA Code for classifying injury, I.E. 10 – Injury, 26 – Illness
HospDay	
HospCost	Cost of hospital days, based on formula in DOD 6055.7
LostTime	Number of LWD
LostPerDay	Figure for LWD, based on computation in DOD 6055.7.
TotalLostTime	Total of hospital and LWD
LostCost	Cost for LWD, based on computation in DOD 6055.7.
LimitLightDuty	Number of limited or light duty days.

LimitLightCost	Cost for light duty days, based on computation in DOD 6055.7.
TotalTimeCost	Total cost of hospital and LWD
NoLostTime	Checked or no LWD (not used)
FirstAid	No definition given
GeneralActivity	Activity the person was performing at the time of mishap.
SpecificActivity	Specific activity personnel was engaged in at the time of the mishap
StateDriversLicence	Checked if personnel injured or involved in the mishap had a license to operate vehicle, machinery, etc
RequiresEyeglasses	No definition given
OnlyDaytimeDriving	No definition given
OtherRestrictionsToLicence	No definition given
StateOfLicence	State were license was issued
ExpirationOfLicence	Date the license expires
Alcohol	Was Alcohol used by person injured or person who caused injury
BAC	The Blood Alcohol Content of the personnel who used Alcohol
Drugs	Checked if legal or illegal drugs were used during the mishap.
LegalDrugs	Prescription or over-the-counter drugs used at the time of the mishap
IllegalDrug	Type of illegal drug that was used at the time of the mishap
FormalTraining	No definition given
FormalTrainingPlace	Any training completed that was related to the mishap.
FormalTrainingDate	Date training was completed.
DICTraining	Any training completed that related to the mishap. I.E. formal training
DICTrainingPlace	Location where person completed the training. I.E. Camp Lejuene, MCB Quantico
DICTrainingDate	Date the person completed the training.
MotorcycleTraining	Any training completed related to mishap. I.E. Formal training
MotorcycleTrainingPlace	Location where person completed training. I.E. Camp Lejuene, MCB Quantico
MotorcycleTrainingDate	Date training was completed
PPE	No definition given
GeneratedId	No definition given

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. S-PLUS CODE FOR ANALYSIS

Classification Tree:

```
treetest<-tree(as.factor(BinRate) ~ factor(COCCodeA)+factor(BaseCodeN)+ Age + PayGrade, data = MC2006Regr, na.action = na.exclude, mincut = 500, minsize = 1000, mindev = 0.01)

summary (treetest)

treetest.cv.m <- cv.tree (treetest, FUN=prune.tree)

min.cut <-treetest.cv.m$size[treetest.cv.m$dev == min(treetest.cv.m$dev)]
min.cut
plot(treetest.cv.m)
title(main="Cross-Validation Plot of CY06 Classification Tree")

treetest.2 <- prune.tree (treetest, best=8)
plot(treetest.2)
plot(treetest.2, type="u")
text(treetest.2, pretty=0)
summary(treetest.2)

post.tree(treetest.2, "CY2006 CART (Pruned by Misclassification Rate)", file=paste(title, "treetest2.ps", sep = ""), digits=.Options$digits - 3, pretty=0, pointsize=12)
```

Generalized Additive Model:

```
*** Generalized Additive Model ***

Call: gam(formula = BinRate ~ s(Age) + s(PayGrade) + factor(COCCodeA) + factor(BaseCodeN), family = binomial(link = logit), data = MC2006Regr, na.action = na.exclude, control = list(epsilon = 0.001, bf.epsilon = 0.001, maxit = 50, bf.maxit = 10, trace = F))

Deviance Residuals:
       Min         1Q     Median         3Q      Max
 -0.4123157 -0.2094875 -0.1321697 -0.09574855 3.730625

(Dispersion Parameter for Binomial family taken to be 1 )

    Null Deviance: 2812.399 on 17469 degrees of freedom
```

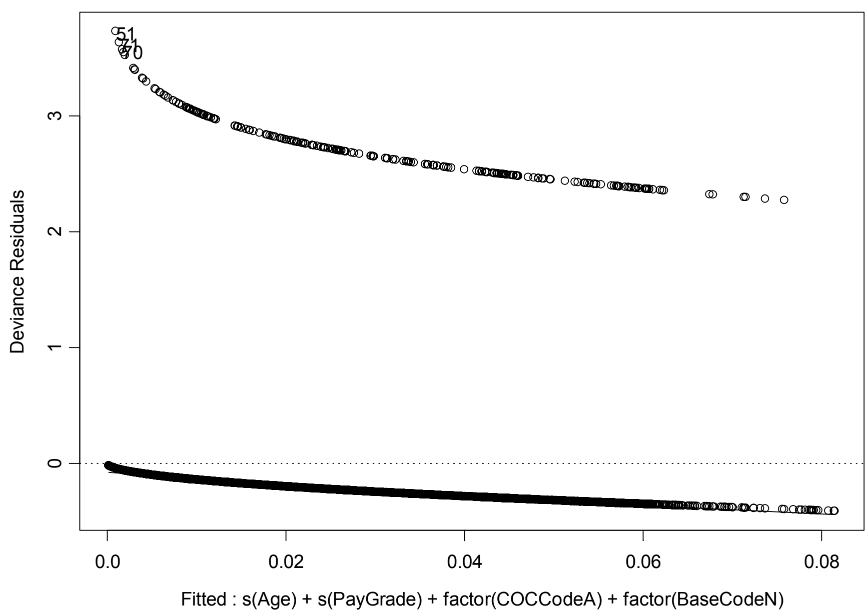
Residual Deviance: 2571.31 on 17448.34 degrees of freedom

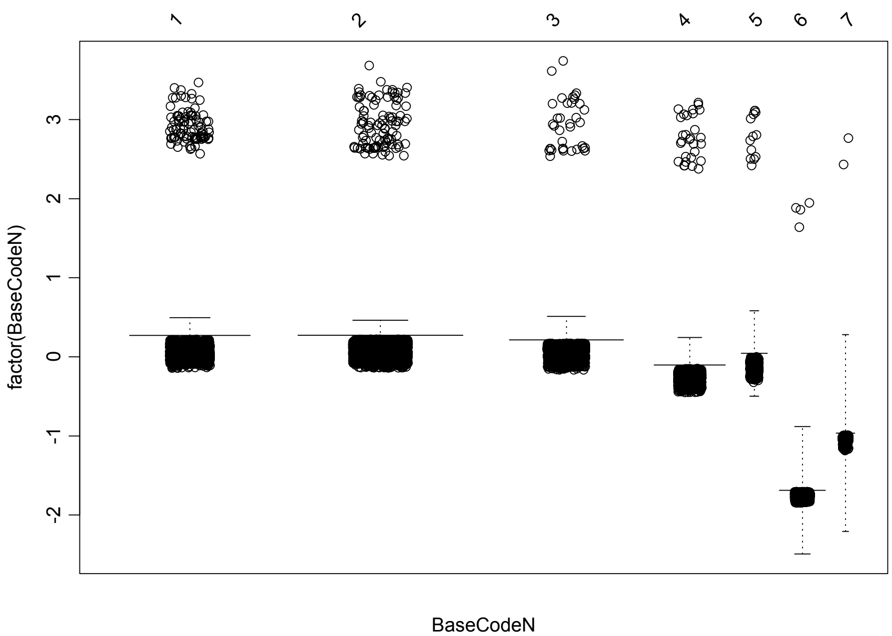
Number of Local Scoring Iterations: 6

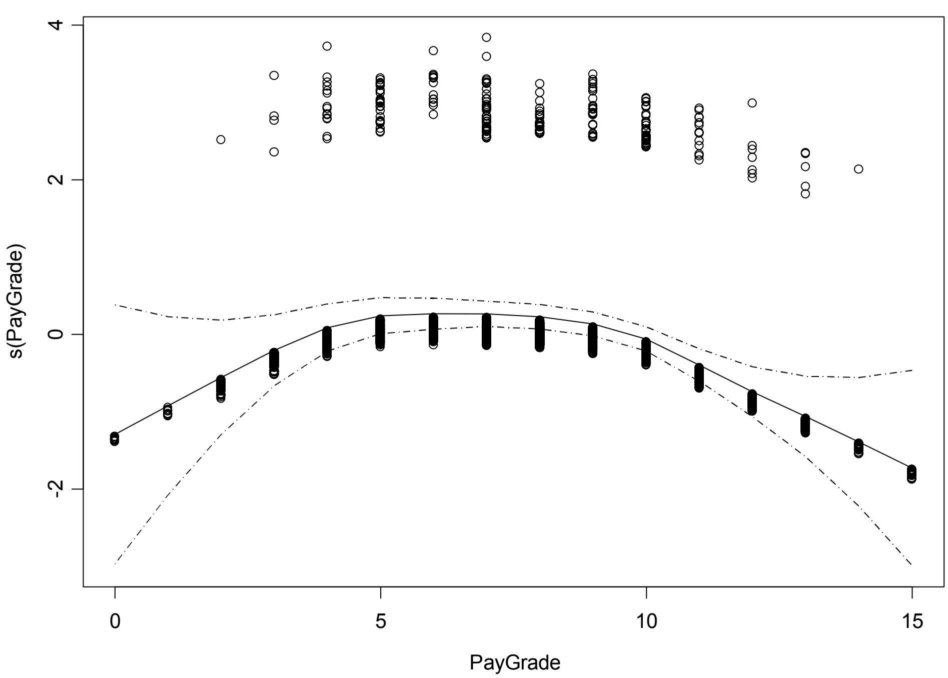
DF for Terms and Chi-squares for Nonparametric Effects

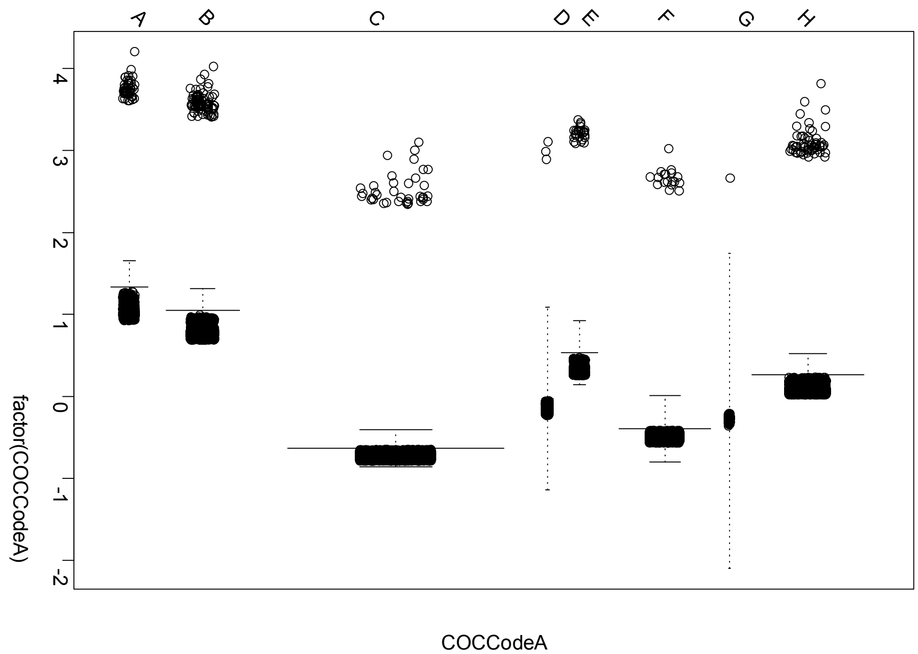
	Df	Npar Df	Npar Chisq	P(Chi)
(Intercept)	1			
s(Age)	1	2.9	4.24865	0.2208618
s(PayGrade)	1	2.8	19.82089	0.0001411
factor(COCCodeA)	7			
factor(BaseCodeN)	6			

The following are the residual plots from the generalized additive model:

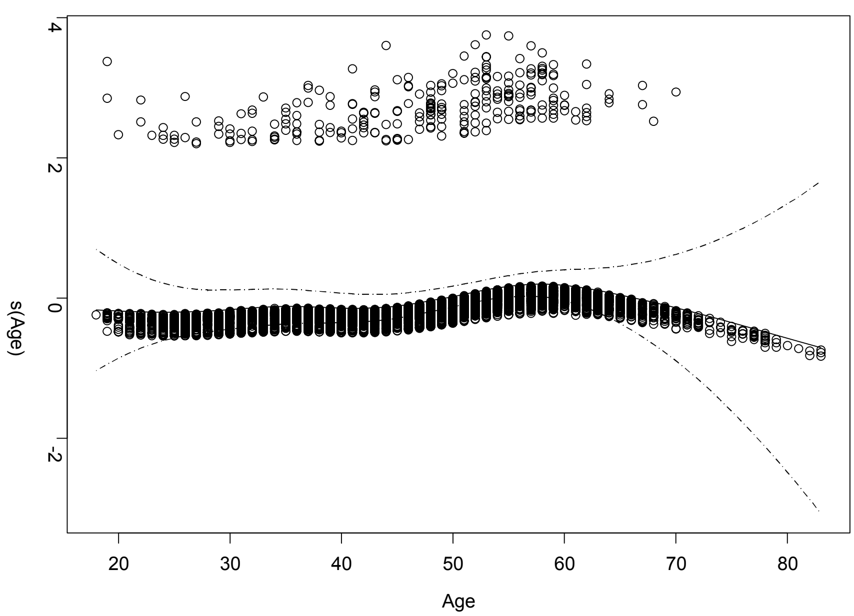



1
2
3
4
5
6
7
3
2
1
0
-1
-2
factor(BaseCodeN)
BaseCodeN


4
2
0
-2
s(PayGrade)
0
5
10
15
PayGrade



A
B
C
D
E
F
G
H
factor(COCCodeA)
COCCodeA



s(Age)
Age

APPENDIX D. COC DESIGNATION

COCCode		
JobTitle	**COC**	**Code**
Fire Protection and Prevention	00081	A
Police	00083	A
Game Law Enforcement	01812	A
Fire Protection Engineering	00804	A
Security Administration	00080	A
Security Clerical and Assistance	00086	A
Security Guard	00085	A
United States Marshal	00082	A
Heavy Mobile Equipment Mechanic	58003	B
Maintenance Mechanic	47049	B
Electronics Mechanic	26004	B
Electronic Measurement Equipment Mechanic	26002	B
Ordnance Equipment Mechanic	66041	B
Heating and Boiler Plant Equipment Mechanic	53009	B
Air Conditioning Equipment Mechanic	53006	B
Industrial Equipment Mechanic	53052	B
Automotive Mechanic	58023	B
General Equipment Mechanic	47037	B
Mobile Equipment Metal Mechanic	38009	B
Electronic Integrated Systems Mechanic	26010	B
Mobile Equipment Servicing	58006	B
Railroad Repairing	35046	B
Sheet Metal Mechanic	38006	B
Digital Computer Mechanic	26008	B
Electromotive Equipment Mechanic	58076	B
Electronic Industrial Controls Mechanic	26006	B
Fuel Distribution System Mechanical	42055	B
Instrument Mechanic	33059	B
Marine Machinery Mechanic	53034	B
Miscellaneous Transportation/Mobile Equipment Maintenance	58001	B
Pneudraulic Systems Mechanic	82055	B
Powered Support Systems Mechanic	53078	B
Small Engine Mechanic	86010	B
Telecommunications Mechanic	25002	B
Miscellaneous Clerk and Assistant	00303	C
General Business and Industry	01101	C
Production Control	01152	C
Human Resources Assistance	00203	C

COCCode		
JobTitle	COC	Code
Quality Assurance	01910	C
Equal Employment Opportunity	00260	C
Logistics Management	00346	C
Miscellaneous Administration and Program	00301	C
General Legal and Kindred Administration	00901	C
Secretary	00318	C
Contracting	01102	C
Supply Program Management	02003	C
Housing Management	01173	C
Accounting Technician	00525	C
Management and Program Clerical and Assistance	00344	C
Office Automation Clerical and Assistance	00326	C
Cash Processing	00530	C
Financial Administration and Program	00501	C
Human Resources Management	00201	C
Inventory Management	02010	C
Financial Clerical and Assistance	00503	C
Accounting	00510	C
Administration and Office Support Student Trainee	00399	C
Administrative Officer	00341	C
Budget Analysis	00560	C
Budget Clerical and Assistance	00561	C
Building Management	01176	C
Business and Industry Student Trainee	01199	C
Civilian Pay	00544	C
Clerk-Typist	00322	C
Commissary Management	01144	C
Computer Clerk and Assistant	00335	C
Data Transcriber	00356	C
Distribution Facilities and Storage Management	02030	C
Editorial Assistance	01087	C
Facility Management	01640	C
Financial Management	00505	C
Financial Management Student Trainee	00599	C
General Attorney	00905	C
General Supply	02001	C
Hearings and Appeals	00930	C
Industrial Property Management	01103	C
Legal Assistance	00986	C
Mail and File	00305	C
Management and Program Analysis	00343	C

COCCode		
JobTitle	**COC**	**Code**
Manpower Development	00142	C
Military Pay	00545	C
Paralegal Specialist	00950	C
Printing Management	01654	C
Procurement Clerical and Technician	01106	C
Program Management	00340	C
Sales Store Clerical	02091	C
Statistical Assistant	01531	C
Support Services Administration	00342	C
Tax Specialist	00526	C
Technical Writing and Editing	01083	C
Traffic Management	02130	C
Transportation Clerk and Assistant	02102	C
Transportation Operations	02150	C
Transportation Specialist	02101	C
Writing and Editing	01082	C
Safety and Occupational Health Management	00018	D
Biological Science Technician	00404	D
Biomedical Engineering	00858	D
Dental Assistant	00681	D
Dietitian and Nutritionist	00630	D
General Biological Science	00401	D
General Health Science	00601	D
Health Aid and Technician	00640	D
Health Physics	01306	D
Health System Specialist	00671	D
Medical Records Administration	00669	D
Medical Records Technician	00675	D
Medical Support Assistance	00679	D
Nurse	00610	D
Public Health Educator	01725	D
Motor Vehicle Operator	57003	E
Sandblasting	54023	E
Boiler Plant Operator	54002	E
Tractor Operator	57005	E
Crane Operating	57025	E
Fork Lift Operating	57004	E
Water Treatment Plant Operator	54009	E
Fuel Distribution System Operator	54013	E
Machining	34014	E
Utility Systems Repairer-Operator	47042	E

COCCode		
JobTitle	**COC**	**Code**
Miscellaneous Industrial Equipment Operation	54001	E
Welding	37003	E
Engineering Equipment Operating	57016	E
Braking-Switching and Conducting	57036	E
Electric Power Controller	54007	E
Equipment Operator	00350	E
Equipment Specialist	01670	E
Equipment, Facilities and Services Assistance	01603	E
Laundry Machine Operating	73005	E
Miscellaneous Transportation/Mobile Equipment Operation	57001	E
Small Craft Operating	57086	E
Telephone Operating	00382	E
Utility Systems Operator	54006	E
Wastewater Treatment Plant Operator	54008	E
Information Technology Management	02210	F
Engineering Technician	00802	F
Electrician (High Voltage)	28010	F
Telecommunications Processing	00390	F
Electrical Equipment Repairer	28054	F
Electrician	28005	F
Materials Examining and Identifying	69012	F
Communications Line Installing and Repairing	25008	F
General Engineering	00801	F
Miscellaneous Electrical Installation and Maintenance	28001	F
Chemical Engineering	00893	F
Chemistry	01320	F
Civil Engineering	00810	F
Communications Clerical	00394	F
Computer Engineering	00854	F
Computer Operation	00332	F
Computer Science	01550	F
Electrical Engineering	00850	F
Electronics Engineering	00855	F
Electronics Technician	00856	F
Engineering and Architecture Student Trainee	00899	F
Engineering Drafting	00818	F
General Telecommunications	00392	F
Industrial Engineering	00896	F
Industrial Engineering Technician	00895	F
Intelligence	00132	F
Locomotive Engineering	57037	F

COCCode		
JobTitle	**COC**	**Code**
Materials Engineering	00806	F
Mathematical Statistician	01529	F
Mathematics	01520	F
Mechanical Engineering	00830	F
Miscellaneous Wire Communications Equipment Installation and	25001	F
Nuclear Engineering	00840	F
Operations Research	01515	F
Physical Science Technician	01311	F
Physics	01310	F
Psychology	00180	F
Public Affairs	01035	F
Purchasing	01105	F
Safety Engineering	00803	F
Safety Technician	00019	F
Social Science	00101	F
Social Science Aid and Technician	00102	F
Social Services	00187	F
Social Services Aid and Assistant	00186	F
Social Work	00185	F
Technical Information Services	01412	F
Telecommunications	00391	F
Welding Engineering	00894	F
Wire Communications Cable Splicing	25004	F
Aircraft Servicing	88062	G
Air Traffic Control	02152	G
Aircraft Freight Loading	69068	G
Aircraft Mechanic	88052	G
Airfield Clearing Equipment Operating	57067	G
Aviation Safety	01825	G
Recreation Specialist	00188	H
Laboring	35002	H
Supply Clerical and Technician	02005	H
Coal Handling	69003	H
Painting	41002	H
Tools and Parts Attending	69004	H
Electroplating	37011	H
Materials Handler	69007	H
Cartographic Technician	01371	H
Plumbing	42006	H
Pipefitting	42004	H
Education and Training Technician	01702	H

COCCode		
JobTitle	**COC**	**Code**
Materials Expediting	69010	H
Tile Setting	36004	H
Environmental Protection Specialist	00028	H
Artillery Repairing	66005	H
Equipment Cleaning	70009	H
Miscellaneous Laundry, Dry Cleaning, and Pressing	73001	H
Locksmithing	48004	H
Dispatching	02151	H
Sign Painting	41004	H
General Education and Training	01701	H
Recreation Aid and Assistant	00189	H
Library Technician	01411	H
Miscellaneous Warehousing and Stock Handling	69001	H
Packing	70002	H
Small-Arms Repairing	66010	H
Transportation Loss and Damage Claims Examining	02135	H
Agronomy	00471	H
Animal Caretaking	50048	H
Archeology	00193	H
Architecture	00808	H
Archives Technician	01421	H
Archivist	01420	H
Audiovisual Production	01071	H
Auditing	00511	H
Blocking and Bracing	46002	H
Botany	00430	H
Bowling Equipment Repairing	48019	H
Cargo Scheduling	02144	H
Carpenter	46007	H
Cement Finishing	36002	H
Clothing Design	00062	H
Community Planning	00020	H
Community Planning Technician	00021	H
Compliance Inspection and Support	01802	H
Construction Control	00809	H
Contact Representative	00962	H
Cook	74004	H
Correctional Institution Administration	00006	H
Custodial Worker	35066	H
Domestic Appliance Repairing	48055	H
Ecology	00408	H

COCCode		
JobTitle	COC	Code
Education and Vocational Training	01710	H
Education Program	01720	H
Education Services	01740	H
Environmental Engineering	00819	H
Environmental Protection Assistant	00029	H
Equal Opportunity Assistance	00361	H
Exhibits Making/Modeling	47015	H
Exhibits Specialist	01010	H
Fabric Working	31005	H
Floor Covering Installing	36009	H
Food Service Worker	74008	H
Forestry	00460	H
Forestry Technician	00462	H
Freight Rate	02131	H
Gardening	50003	H
General Arts and Information	01001	H
General Facilities and Equipment	01601	H
General Inspection, Investigation, and Compliance	01801	H
General Investigating	01810	H
General Physical Science	01301	H
General Student Trainee	00099	H
Geography	00150	H
Geology	01350	H
Glazing	36011	H
Hazardous Waste Disposer	69013	H
History	00170	H
Human Resources Management Student Trainee	00299	H
Hydrology	01315	H
Illustrating	01020	H
Industrial Hygiene	00690	H
Industrial Specialist	01150	H
Information and Arts Student Trainee	01099	H
Information Receptionist	00304	H
Instructional Systems	01750	H
Insulating	36010	H
Laundry and Dry Cleaning Equipment Repairing	53017	H
Laundry and Dry Cleaning Plant Management	01658	H
Laundry Working	73004	H
Legal Instruments Examining	00963	H
Librarian	01410	H
Masonry	36003	H

COCCode		
JobTitle	**COC**	**Code**
Messenger	00302	H
Metal Tank and Radiator Repairing	38058	H
Metalizing	37007	H
Miscellaneous Ammunition, Explosives, and Toxic Materials Wo	65001	H
Miscellaneous Armament Work	66001	H
Miscellaneous Electronic Equipment Installation and Maintena	26001	H
Miscellaneous General Equipment Maintenance	48001	H
Miscellaneous General Maintenance and Operations Work	47001	H
Miscellaneous Industrial Equipment Maintenance	53001	H
Miscellaneous Machine Tool Work	34001	H
Miscellaneous Metal Processing	37001	H
Miscellaneous Metal Work	38001	H
Miscellaneous Packing and Processing	70001	H
Miscellaneous Painting and Paperhanging	41001	H
Miscellaneous Plant and Animal Work	50001	H
Miscellaneous Plumbing and Pipefitting	42001	H
Miscellaneous Woodwork	46001	H
Motor Carrier Safety	02123	H
Museum Curator	01015	H
Museum Specialist and Technician	01016	H
Naval Architecture	00871	H
Optical Instrument Repairing	33006	H
Packaging	02032	H
Passport and Visa Examining	00967	H
Pest Controller	50026	H
Photography	01060	H
Plastering	36005	H
Preservation Service	70006	H
Pressing	73006	H
Property Disposal	01104	H
Property Disposal Clerical and Technician	01107	H
Rangeland Management	00454	H
Realty	01170	H
Rigging	52010	H
Roofing	36006	H
Soil Conservation	00457	H
Soil Conservation Technician	00458	H
Sports Specialist	00030	H
Store Working	69014	H
Test Range Tracking	52035	H
Toolmaking	34016	H

COCCode		
JobTitle	**COC**	**Code**
Training Instruction	01712	H
Tree Trimming and Removing	50042	H
Upholstering	31006	H
Visual Information	01084	H
Voucher Examining	00540	H
Wildlife Biology	00486	H
Wood Crafting	46005	H
Wood Worker	46004	H

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bowes, M., “An Analysis of DoN Mishap Rates,” (CRM D0009828.A2/Final). Alexandria, VA: The Center for Naval Analyses, 2004.

Bowes, M., “Workers’ Compensation: The Case for Improved DoN Management,” (CRM D0008209.A2/Final). Alexandria, VA: The Center for Naval Analyses, 2003.

Conatser, D., “Forecasting U.S. Marine Corps Reenlistments By Military Occupational Specialty And Grade,” Master’s Thesis, Naval Postgraduate School, Monterey, California, 2006.

Defense Manpower Data Center. <https://www.dmdc.osd.mil/pprofile/owa/intro.intro>, 11 August 2007.

Department of Defense, Civilian Personnel Management Service. <http://www.cpms.osd.mil/hrbits/userguide/dcpds_userguide.aspx>, 11 August 2007.

Devore, J.L., “Probability and Statistics: For Engineering and the Sciences,” Thompson Brooks/Cole, 2004.

Employment Standards Administration: U.S. Department of Labor, <www.dol.gov/esa>, 11 August 2007.

Hair, J.F., Anderson, R.E. and Tatham, R.L., “Multivariate Data Analysis,” Macmillan Publishing Company, 1987.

Hand, D., Mannila, H. & Smyth, P., “Principles of Data Mining,” The MIT Press, 2001.

Hinson, W.B., “A Statistical Analysis of Individual Success After Successful Completion of Defense Language Institute Foreign Language Center Training,” Master’s Thesis, Naval Postgraduate School, Monterey, California, 2005.

Integrated Publishing, “Safety Policy,” <http://tpub.com/content/administration/14167/css/14167_13.htm>, 11 August 2007.

McWilliams, N., “Analysis of Lost Workday Rate For Marine Corps Logistics Base Barstow, CA.” Norfolk, VA: Naval Safety Center, 2006.

Montgomery, D.C., Peck, E.A. and Vining, G.G., “Introduction to Linear Regression Analysis,” A Wiley-Interscience Publication, 2001.

Naval Audit Service, “Reducing Lost Work Time Due to On-the-Job Injuries at Navy and Marine Corps Commands.” (N2004-0034). Washington Navy Yard, DC, 2004.

Naval Safety Center, <http://www.safetycenter.navy.mil/services/mission.htm>, 11 August 2007.

Reed, J.H., “Question about ERP.” Email to author. 4 April 2007.

Rice, J.A., “Mathematical Statistics and Data Analysis,” Wadsworth, Inc., 1988.

Rockett, T., “SEASHORE SQL Scripts.” E-mail to author. 9 July 2007.

Standard Labor Data Collection and Distribution Application, <http://www.sldcada.disa.mil/sldevent/Index.do>, 11 August 2007.

United States Department of Justice, <http://www.usdoj.gov/oip/privstat.htm>, 11 August, 2007.

Weiss, N.A., “Introductory Statistics,” Addison-Wesley Publishing Company, Inc., 1995.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. United States Marine Corps, Safety Division
   Headquarters, Marine Corps
   Quantico, Virginia

4. Naval Safety Center
   United States Navy
   Norfolk, Virginia

5. Professor Lyn Whitaker
   Naval Postgraduate School
   Monterey, California

6. Commanding General, Training and Education Command
   MCCDC, Code C46
   Quantico, Virginia

7. Director, Marine Corps Research Center
   MCCDC, Code C40RC
   Quantico, Virginia

8. Marine Corps Tactical Systems Support Activity (Attn: Operations Officer)
   Camp Pendleton, California

9. Director, Operations Analysis Division
   Code C19, MCCDC
   Quantico, Virginia

10. Marine Corps Representative
    Naval Postgraduate School
    Monterey, California

11. MCCDC OAD Liaison to Operations Research Department
    Naval Postgraduate School
    Monterey, California